

Constraint-Based Phonology

Steven Bird

Ph.D.
University of Edinburgh
1990



Abstract

This thesis presents the results of a study in the application of logic to phonology, the subfield of linguistics concerned with the 'sound structure' of the world's languages. The logical framework is classical first order predicate calculus with a model-theoretic semantics. Existing proposals in temporal logic (van Benthem) and feature logic (Johnson) are combined in the treatment of temporal and hierarchical organization. Phonological 'representations' are linguistic descriptions couched in a formal language. The set of utterance tokens forms the class of intended models. Some links with the sign-based view of grammatical and lexical organization are explored, with a view to ultimately supplementing sign-based linguistic theories (such as Head-Driven Phrase Structure Grammar) with phonological information.

A model of feature organization based on phonological argumentation (following Sagey) and phonetic argumentation (following Browman & Goldstein) is proposed as an exemplification of the approach. The model achieves a clear distinction between articulatory and acoustic classificatory properties, lending clarity to the debate about the function of the so-called manner features, and giving content to some recent calls for a non-segmental phonology.

Arising from this logical approach is a new computational metaphor for phonology, namely constraint-satisfaction. Linguistic generalizations may be stated in the declarative style, liberated from concerns about their procedural implementation in performance tasks such as generation and recognition. A working constraint-solver which interfaces to Prolog is described.

Declaration

I declare that this thesis has been composed by me and that the work reported here is original except where acknowledged otherwise.

Steven Bird

28th September 1990.

Acknowledgements

This work owes a considerable debt to Ewan Klein, friend and teacher, who has enthusiastically supported and encouraged me throughout. I feel privileged to have been guided by someone with his vision and incisiveness. His detailed comments on numerous drafts of this thesis have invariably revealed the flaws in my reasoning and shown a way forward. Robin Cooper and Bob Ladd have also acted as supervisors; the breadth and depth of their interests have been inspirational.

Roland Sussex and John Upton fostered my nascent interests in linguistics and mathematics respectively, and convinced me of the potential for the fruitful interaction of the two domains. I hesitate to think where I would have been without them. Heather Bowe first introduced me to phonology, the richness and complexity of which intrigues me today as much as it did then. Jim Scobbie, Michael Broe, Mike Reape, Jo Calder and Pete Whitelock have been soul-mates over the past three years, sharing my frustrations and fascinations. Others who provided excellent teaching, invaluable suggestions and stimulating debate along the way include Diana Archangeli, Patrick Blackburn, Chris Brew, Ted Briscoe, Nick Clements, Jennifer Cole, John Coleman, John Goldsmith, Franz Guenther, James Harland, Bruce Hayes, Jack Hoeksema, Lex Holt, Mark Johnson, Andras Kórnai, Marcus Kracht, John McCarthy, Chris Mellish, Bill Poser, Alan Prince, Doug Pulleyblank, Geoff Pullum, Bill Rounds, Jerry Seligman, Gary Simons, Nigel Vincent and Shelley Waksler. Thanks to you all for helping me distinguish the wheat from the chaff. Robin Cooper, Dafydd Gibbon, Dick Oehrle, Henry Thompson, Richard Weise and Pete Whitelock provided invaluable feedback on the final draft. Without their painstaking efforts few readers would have been able to make sense of what I have written.

Edinburgh has been a truly wonderful place to undertake a Ph.D. The international and interdisciplinary nature of the Centre for Cognitive Science has ensured a steady flow of interesting people and new ideas. This is one of the few places in the world that a computer scientist can do a Ph.D. in phonology and have plenty of people to talk to.

The Linguistic Society of America sponsored my attendance at the Linguistic Institute in Tucson, Arizona in 1989. I am grateful for the opportunity of presenting this work at the Second European Summer School on Logic, Language and Information in Leuven, Belgium in 1990, and to those who attended the course and provided feedback. I am thankful to the University of Edinburgh and the Victoria League for their generous financial support.

Greater debts are nearer to home. My christian friends have been an extended family to me, providing continual encouragement and helping me retain my sanity in the final stages of writing up. My parents, by their example, taught me perseverance and dedication to the task, for which I am eternally grateful. The richest blessing on my life has been Kay. This thesis is a testimony to her strength, love and support.

Table of Contents

1	Introduction.....	1
1.1	Formal Phonology — A Model-Theoretic View.....	2
1.2	Monostratal Phonology.....	8
1.3	Prosodic Phonology.....	15
1.4	Sign-Based Phonology.....	20
1.5	Overview of Thesis.....	24
2	A Logical Foundation for Phonology.....	26
2.1	Sorts.....	27
2.2	Hierarchical Organization in Phonological Structures.....	31
2.3	Temporal Organization in Phonological Structures.....	38
2.4	The Interaction of Hierarchical and Temporal Structure.....	43
2.5	Another Graphical Notation: The Feature Matrix.....	45
2.6	Some Abbreviatory Conventions.....	49
2.7	Rules and Principles.....	51
2.8	The Language \mathcal{L}	54
2.9	Combining Descriptions.....	56
2.10	Conclusion.....	58
3	A Theory of Phonological Structure.....	60
3.1	The Evidence for Hierarchical Organization.....	62
3.2	An Articulatory Model.....	77
3.3	Prosodic Structure.....	87
3.4	Formalizing the Theory.....	99
3.5	Three Kinds of Phonological Generalization.....	102
3.6	Conclusion.....	105
4	Implementation.....	106
4.1	Representation.....	108
4.2	An Example.....	111
4.3	Conclusion.....	113
5	Conclusion.....	116
	References.....	121
	Appendix.....	133
	Appendix A: Extract from Phonological Events.....	133
	Appendix B: Feature Structures and Indices.....	141

Chapter 1

Introduction

The fundamental goal of linguistics is to discover a systematic relation between the domain of spoken utterances and the domain of possible meanings. This relation can only be systematic insofar as the domains it relates possess internal structure. The present study is devoted to an elucidation of the structure and organization of PHONOLOGY, that area of research devoted to the study of the linguistically significant aspects of speech.

CONSTRAINT-BASED PHONOLOGY is an intensional theory of phonology tailored to the sign-based approach to linguistic description. It differs from other phonological theories in its use of the description/object (or type/token) distinction and in its rejection of the representation/rule distinction. The purpose of this chapter is to elaborate the four cornerstones of Constraint-Based Phonology: (i) a formal foundation, (ii) monostratal analysis, (iii) prosodic structure and (iv) sign-based grammar. Chapter two picks up on the discussion of point (i) and contains a detailed proposal for a phonological formalism. Chapter three takes up some of the issues raised in the discussion of point (iii). Chapter four presents a computational interpretation of the formalism, and chapter five makes some general observations and concludes the present study.

1.1 Formal Phonology — A Model-Theoretic View

The descriptive apparatus of SPE (Chomsky & Halle 1968) was sufficiently formal that the generative capacity of a grammar could be established (Johnson 1972). Unfortunately, few of the frameworks developed since then have achieved this basic level of explicitness. (See Bird & Ladd (*forthcoming*) for a critique of autosegmental phonology along these lines.) Although few would dispute the centrality of formalism to phonology, the very notion FORMAL often seems enigmatic (cf. Pullum 1989). If we begin by accepting the distinction between a grammar and the language in which it is expressed, it is immediately obvious that talk of formal grammar presumes the existence of a formal language for linguistic description. While it has been generally assumed that prose does not qualify as a formal language (perhaps simply because natural language is “ill adapted for talking about itself” Firth 1948/1957:121–2) it seems less clear exactly what is required of artificial languages (such as the graphical notations of phonology) before they can be called formal. What is clear, however, is that employing a (possibly graphical) description language which is defined using prose is not necessarily any more formal than using prose directly to describe the phenomena at hand.

The most fundamental requirement for a formal language is that its membership is well-defined. We must be able to decide, for a given expression, whether or not it is a member of the language. This is normally achieved by using a formal syntax, a recursive definition of well-formedness (Chomsky 1959). The notations adopted in phonology appear to demand an adaptation of this requirement to the graphical domain: “Phonologists characteristically encode their representations graphically, as lines on a page. In principle, this is a good idea, since it usually makes the representations easier to visualize and manipulate. But it is important to have in mind a clear notion of what graphic formalisms are intended to stand for” (Hayes 1990:39). The goal of the present study is more humble, however. Rather than defining the syntax of a graphical language, I employ a formal language (in the non-graphical sense of the term) and adopt certain conventions for depicting logical

descriptions (cf. Johnson 1988:28).

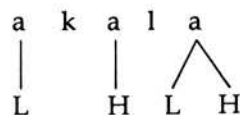
Aside from the need for a formal syntax there is a second requirement of a formal language that has to do with the relationship between each expression of that language and the state(s) of affairs in the world that the expression describes. Indeed, this concern applies to science in general:

...when we construct a particular science it is not sufficient to specify the levels and state their theorems: a science is not just a framework for us to admire, it is essentially a tool. A science is constructed with a particular purpose in mind, to provide a critical and selective description. In order to employ a science in this way *we must have some means of interpreting its descriptions*; and we can do this by appeal to meaning. (Dixon 1963:36, emphasis added)

In the present context, this is achieved by having a formal semantics for the description language. The meaning of an expression e of the language \mathcal{L} will be defined as the set of utterances e describes. On this view there are in fact two languages, one describing (i.e. \mathcal{L}) and one being described (say \mathcal{U}), and a relation between them specifying which members of \mathcal{L} describe which members of \mathcal{U} .

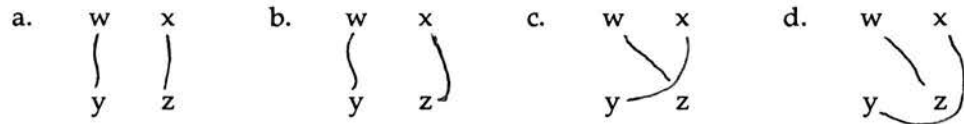
As an illustration of what is meant by syntax and semantics in the context of a language for phonological description, we turn to what is still the most detailed attempt to provide a formal foundation for autosegmental phonology, namely Goldsmith's landmark dissertation (1976). The first chapter of Goldsmith's thesis includes an informal definition of a graphical language. Its main features are sequences of feature matrices (or abbreviatory characters), arrayed on two LEVELS (or TIERS) and lines drawn between pairs of matrices or characters on different levels, called ASSOCIATION. This notation is illustrated in (1-1) for the word *àkálǎ*. Here, the tonal information expressed orthographically using the diacritic marks ` (low tone), ´ (high tone) and ~ (rising tone) is separated out using the symbols L and H . The former connection between vowels and diacritics is recorded using association lines, and the rising tone is analysed as a sequence of a low tone followed by a high tone.

(1-1)



Let us define \mathcal{G} to be the space of all possible diagrams, broadly construed to include all potential arrangements of ink on a page, meaningful or not. Goldsmith's informal definition circumscribes the space of autosegmental diagrams \mathcal{A} from the rest of \mathcal{G} . Since his definition is only informal, the precise extent of \mathcal{A} is unclear. For example, if the lines do not have to be perfectly straight then how curved can they be?

(1-2)



I shall later argue that the only restriction needed is for all diagrams to correspond to well-formed expressions of \mathcal{L} . In this way, we can accept (1-2a,b) but rule out (1-2c,d) as ill-formed.

A further refinement is provided by a well-formedness condition (Goldsmith 1976:27), restated in (1-3)¹.

(1-3) **Well-formedness Condition**

- a. Every matrix or character must be associated to another matrix or character on the other tier.
- b. Association lines do not cross.

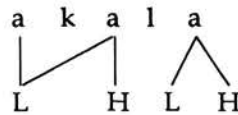
The condition in (1-3b) is known as the no-crossing constraint. The well-formedness condition is a constraint on graphical representations. In effect, it divides the set of autosegmental diagrams \mathcal{A} into the well-formed diagrams \mathcal{A}_w and the ill-formed diagrams \mathcal{A}_i . Could the well-formedness condition be the formal syntax for the graphical language of autosegmental phonology we are seeking? The simple answer is no. The standard generative understanding of well-formedness is that phonological rules are not restricted to apply only to elements of \mathcal{A}_w but are free to apply to all of \mathcal{A} . A *derivation*² is ultimately just a composite function having \mathcal{A} as its domain and \mathcal{A}_w as its range (Poser 1982:124). Only the final stage of a derivation must produce a well-formed representation. We can observe two notions of well-formedness here: the first defines the kinds of objects a phonological rule can operate on, and the second defines the kinds of objects given a phonetic interpretation.

¹ This well-formedness condition assumes the existence of only two tiers. However, a multi-tiered representation—a tree of tiers—would require this condition to be met by all pairs of associable tiers.

² A derivation is a process by which a particular lexical form is converted into a surface form.

A consequence of this duality has been uncertainty in the literature about the formal status of the no crossing constraint, persisting through to (Goldsmith 1990). One interpretation is that the no crossing constraint blocks the application of a rule which would result in crossing lines (1990:30). Another view is that the no crossing constraint *repairs* ill-formed representations by causing the deletion of the line crossed by the new line (1990:47,79)³. For example, consider the effect of an (iterative) rule which spreads the initial tone of a word, applied to (1-1). On the first application, a line from the *L* to the second *a* is added:

(1-4)



Under the 'blocking' view of the no-crossing constraint, the rule could not re-apply, since linking the *L* to the third *a* would cause a line crossing. Under the 'repair' view, there are two further steps. In (1-5a) a line which violates the no-crossing constraint is added (the asterisk indicates this ill-formedness), and in (1-5b), the old line which this new line crossed is deleted, leaving the first *H* tone unassociated. (Presumably a further derivational step would associate this *H* tone, perhaps to the final *a*.)

(1-5)



Therefore, the different interpretations of the no-crossing constraint are significant. Because derivations are understood to be deterministic, both interpretations cannot co-exist. The reason for this ambiguity is the informal connection between the well-formedness condition and a derivation it is understood to license.

There is a radically different notion of well-formedness which comes from logic and

³ This ambiguity between blocking versus repair views has arisen in other areas of phonology, most notably in connection with the 'Obligatory Contour Principle' (OCP), which requires that two adjacent tier elements must not be identical. Goldsmith (1976:36), citing Leben's (1973) original definition of the OCP, clearly views it as a repair strategy which alters a representation containing adjacent identical elements by fusing them into one. However, McCarthy (1986:222) has advocated the opposite view, where the OCP simply blocks a derivation.

which has been adopted in some quarters of linguistics, which may be summarized as follows:

The Well-Formedness Constraint: Each syntactic rule operates on well-formed expressions of specified categories to produce a well-formed expression of a specified category. (Partee 1979:276)

If Goldsmith's language were recast in such a system then it would be necessary to pick which of \mathcal{A} and \mathcal{A}_w constitutes the well-formed expressions. If we employed his algebraic structures (1976:28–30) as the formal semantics for this language the resulting system would either be *unsound* or *incomplete* relative to the class of intended models. If the set of well-formed expressions is \mathcal{A} then most well-formed expressions do not have an interpretation (unsoundness). If the set of well-formed expressions is \mathcal{A}_w then some expressions which have an interpretation are ill-formed (incompleteness; a serious flaw which is relegated to footnotes 5 and 6, 1976:55). If a logic is not both sound and complete, the connection between its expressions and the objects those expressions describe is seriously impaired. Either phonological representations can describe non-existent utterances or some utterances cannot be represented. This state of affairs is highly undesirable given the stated commitment of the generative enterprise to empirical matters (i.e. “descriptive adequacy”). Indeed, it could be argued that proposals for the adoption of descriptive devices having implicit and impaired syntax and semantics do not fall within the scope of the generative enterprise (cf. Gazdar et al. 1985:6).

An approach which makes this syntax/semantics distinction can be viewed from a MODEL-THEORETIC perspective. A feature of the model-theoretic view is the way salient aspects of the domain being investigated are recapitulated in the description language. An utterance is a coordinated set of movements by the articulators in the vocal tract. As such, an utterance is an event which is situated in space and time. Two of our most basic intuitions about actions in time are that (i) it is possible for independent actions to be simultaneous, and (ii) it is possible for an action to be performed at different times. These two intuitions are central to the descriptive framework of autosegmental phonology. First, if two autosegments are associated then the corresponding physical gestures are required to be coarticulated (Goldsmith

1976:16). Second, if the same autosegment appears twice in a representation, then there are two instances of the corresponding physical gestures. This explicitness about the role played by spatio-temporal structure in phonology was clearly evident in (Goldsmith 1976) yet has been downplayed since (Goldsmith 1990). More recently this connection has been made explicit in work on 'laboratory phonology' (e.g. Browman & Goldstein 1989, esp. p.211).

The model-theoretic view of the phonology-phonetics interface differs markedly from the traditional view of generative phonology. Keating summarizes the traditional view as follows:

The SPE model represents lexical items as matrices of binary-valued phonetic features; each row is a feature, and each column a segment. Phonological rules may change the values of features, or may add or delete segments ... By contrast, phonetic rules *convert* the binary values into quantitative values along continuous phonetic scales ... A further universal phonetic component ... will *convert* these scalar values into a representation of articulations that are continuous in time. (Keating 1984:286–287, *emphasis added*)

It is almost as though abstract theoretical constructs (such as phonological representations) are being successively *converted* into the real objects of the domain itself (i.e. utterances). This view has striking similarities with the view of natural language semantics in 'Government and Binding Theory' (Chomsky 1981), where syntactic structures and meanings (i.e. the 'Logical Form') are related to each other derivationally. It is natural to wonder whether such views are based on a category mistake (Ryle 1949, Dennett 1987:213ff). The model-theoretic view would seem to be more plausible. From this position, an abstract linguistic structure can have a semantic interpretation and a phonetic interpretation. Accordingly, we can view phonological representations as *descriptions* of real-world utterances (cf. Bach 1983).

In this section I have advocated the adoption of a formal approach to autosegmental phonology. This formality is not a purely esoteric concern. It stems from the desire to produce empirically adequate and computationally interpretable grammars, as well as theoretically attractive ones.

1.2 Monostratal Phonology

Beyond empirical adequacy, another central concern of phonology has been the naturalness of an analysis. Within SPE this issue was dealt with via a 'simplicity metric' which was based on such notions as the length of a derivation and the number of symbols required for an analysis. If 'adjacently ordered' rules were structurally similar, then they could be collapsed into a single rule, thereby simplifying the analysis, and conforming to a general view that "grammars where structurally similar rules are adjacent are more 'highly valued' than grammars where such rules are not adjacent" (Kisseberth 1970:292). (Similar concerns are apparent in Kiparsky's (1968:196ff) identification of other relations between adjacently ordered rules.) In the context of his analysis of Yawelmani phonology, Kisseberth (1970) identifies a more pervasive relationship between rules than those already mentioned. Yawelmani words do not contain clusters of more than two consonants, nor do words begin or end with clusters. At first blush, rules of vowel epenthesis and consonant deletion appear unrelated. However, Kisseberth (1970:293) notes that "there are a variety of phonological processes which, it might be said, 'conspire' to yield phonetic representations which contain no word-final clusters and no trilateral clusters". The conclusion is that the significant relationships among rules are not structural but functional, and the ramification for the 'cost' of a grammar is an "evaluation procedure which 'counts' only certain parts of functionally related rules" (303). It is unclear exactly what is meant by this, and although Kisseberth sees his discussion as preliminary, the two decades of research since have not witnessed an elucidation of the relationship between the naturalness of an analysis and the functional unity of the rules it comprises.

Another player in this 'abstractness controversy' was NATURAL GENERATIVE PHONOLOGY. Its main proposals were that there should be no extrinsic rule ordering (Vennemann 1972:110) and that all generalizations should be 'surface true' (Hooper 1976:13, cf. Stanley 1967:421ff), thereby ruling out destructive operations such as deletion and metathesis. The high degree of abstractness such operations permit was argued not to be a part of native speaker competence.

Closely related to the True Generalization Condition is a constraint which has arisen in the extended Montague grammar framework, in such formalisms as Generalized Phrase Structure Grammar (GPSG, Gazdar et al. 1985) and Head-Driven Phrase Structure Grammar (HPSG, Pollard & Sag 1987). It is usually expressed as the requirement that there be only one level of linguistic description; grammars adhering to this constraint are called MONOSTRATAL. For a rule $lhs \rightarrow rhs$ to be a true generalization the arrow must be understood logically: “those objects satisfying the description lhs also satisfy the description rhs ” (cf. the ‘local tree’ interpretation of phrase structure rules in Gazdar et al. 1985). However, two other interpretations of rules are apparent in the literature. The standard interpretation, evident from derivations involving rules like (1–6a), is destructive: “those objects satisfying the description in lhs are modified (minimally) so that they now satisfy the description in rhs ”.

- (1–6) a. $[+voice] \rightarrow [-voice]$
 b. $[] \rightarrow [+voice]$

A third interpretation is based on a notion of defaults (e.g. Stanley 1967). A rule like (1–6b) is interpreted thus: “if an object does not have a value for the property in rhs (in this case, voicing) then it gains the value of that property specified in rhs (here, $[+voice]$)”. A more general case of this rule is where a context is specified on the left hand side, and the rule only applies if (i) that context is met and (ii) it is consistent to assume what appears on the right hand side.

While some default rules have a wide applicability, others are highly restricted. If the general rules apply first, then the restricted ones may not subsequently be able to apply, and so it is frequently assumed that rules with the most restrictive contexts must be applied first, and those with the most general contexts last. This ordering constraint is known as the ELSEWHERE PRINCIPLE (Kiparsky 1973). Ordering restrictions of this kind have been described as ‘non-parochial’ by Pullum (1976) to distinguish them from the highly restricted ordering statements of the form: “the lengthening rule precedes the flapping rule”, which he describes as ‘parochial’. However, it has occasionally been noted (e.g. Karttunen, Koskeniemi & Kaplan 1987:30, Calder 1990) that the ordering produced by the elsewhere principle, while convenient, can

be done away with by enriching the context part of a rule. The context of a 'general rule' is refined by adding the negations of the contexts of all the 'specific rules' with which it potentially clashes. For example, consider the rules in (1-7), where (1-7b) is a default rule.

- (1-7) a. $[+back, +high] \rightarrow [+round]$
 b. $[] \rightarrow [-round]$

The first rule states that high, back segments are also round. The second states that if it is consistent to assume any segment is $[-round]$ then let it be $[-round]$. The Elsewhere Principle orders (1-7b) after (1-7a). However, no ordering statement is required if the context of (1-7b) is enriched in either of the following (logically equivalent) ways, where ' \neg ' represents negation and ' \vee ' represents disjunction.

- (1-8) a. $\neg [+back, +high] \rightarrow [-round]$
 b. $[-back] \vee [-high] \rightarrow [-round]$

Now there is no ordering. Furthermore, there is no need to make use of the default / non-default rule distinction. This does not mean to say that default rules such as (1-7b) should be avoided. On the contrary, the above discussion simply reveals their docile nature; they can be viewed as a useful abbreviatory device.

Criticisms of the monostratal approach

The monostratal view has sometimes been criticized as being insufficiently powerful to enable the expression of observations which demand parochial rule ordering or deletion. As will be shown below, some observations which are claimed to require ordering or deletion actually do not. The metatheoretical claim of natural generative phonology—taken up here—is that observations which cannot be expressed without recourse to ordering or deletion are undesirable and can be encompassed in other ways. One reason to believe this approach might be worth exploring is the surprising successes it has had in natural language syntax where, for example, deletion and movement rules previously assumed to be unavoidable were shown to be unnecessary (Gazdar 1981). The remainder of this section is devoted to a discussion of various objections to the monostratal approach and their rebuttal. Like the objec-

tions, the rebuttals concern explanatory as opposed to empirical questions, and are ultimately subjective arguments in favour of a certain style rather than being formal proofs in any sense.

Most criticisms of Natural Generative Phonology have been based on data involving absolute neutralization. For example, Drescher (1981) cites a familiar example from some dialects of North American English, where the voicing contrast of *t* and *d* in *writer* and *rider* is claimed to be manifested only in the length of the preceding vowel. Two rules, lengthening and flapping, are proposed to account for this behaviour.

(1-9) a. **Lengthening:**

$$a \rightarrow [+long] / - (y) \left[\begin{array}{c} C \\ +voice \end{array} \right]$$

b. **Flapping:**

$$\left\{ \begin{array}{c} t \\ d \end{array} \right\} \rightarrow D / \left[\begin{array}{c} V \\ +stress \end{array} \right] - \left[\begin{array}{c} V \\ -stress \end{array} \right]$$

In order to produce the correct result, the lengthening rule must apply before the flapping rule.

	Underlying form	/raytər/ 'writer'	/raydər/ 'rider'
(1-10)	Lengthening	–	ra:ydər
	Flapping	rayDər	ra:yDər
	Phonetic form	rayDər	ra:yDər

However, both rules are descriptively inadequate. The lengthening rule is incorrect for words like *italic* and *idyllic* where the length distinction does not appear (Selkirk 1982:372). Chen's (1970) cross-linguistic study of the lengthening phenomena indicates that it is more of a physiological matter than a linguistic one⁴. It therefore seems as though both physiological and linguistic factors are at work here. As no clear consensus has yet emerged about the precise nature of this phenomena it would be premature to attempt an analysis here. In any case, the flapping rule is a major simplification from the actual data (e.g. Kahn 1976:56–61). However, even if we presumed that it was descriptively adequate, there is no obvious reason to assume that it is the only correct approach to this data. To say that the distinction—represented

⁴ "... it is well known that a voiceless consonant is articulated with open glottis, whereas a voiced one is made closed glottis. As a result, the intraoral pressure during a voiced consonant closure is relatively low, since the pressure is built up by the air of the mouth cavity *alone*; in the case of a voiceless consonant occlusion, the intraoral pressure is considerably higher, since the volume of air of the mouth *and* lungs is increased. ... the transition from vowel to a voiceless consonant closure would be faster than the transition from vowel to a voiced consonant closure." (Chen 1970:152–3)

orthographically as *t* versus *d*—is phonologically one of voicing is debatable. The ‘voicing’ of English stop consonants is not generally manifested as a phonetic voicing difference at all, since both varieties of stop are phonetically voiceless (except for the flaps; Kahn 1976:41, see also Keating 1984 and Lisker 1986). More generally, Dinnsen (1985:276) has claimed that “every genuine phonological distinction has some phonetic reflex, though not necessarily in the segments which are at the seat of the distinction.” This would indicate that a rule which ‘copies’ an underlying distinction elsewhere (e.g. lengthening) and a rule which removes the underlying distinction (e.g. flapping) are more intimately related than the above analysis would suggest. Analyses based on absolute neutralization—such as Dresher’s and Gussmann’s (1980)—clearly stand in need of revision⁵. The complaint about these analyses, given our concerns about abstractness, is that such exhaustively segmental views obscure the fact that the temporal coordination of articulations is at the heart of the issue here.

Next, consider the case of vowel harmony in the Pasiego dialect of Montañés Spanish. This language has a nine-vowel system consisting of five tense vowels and four lax vowels. Our examples will employ the tense vowels only; these are partitioned according to the following table for the purposes of height harmony.

(1-11)

[+high]	i	u
[-high]	e	o
neutral	a	

In the data tabulated in (1-12), the non-low vowels in the verb roots must agree in height with the stressed vowel in the suffix (McCarthy 1984). Low vowels are transparent to this process. The verb stems are given using the vowels E and O which represent the classes {i, e} and {o, u} respectively.

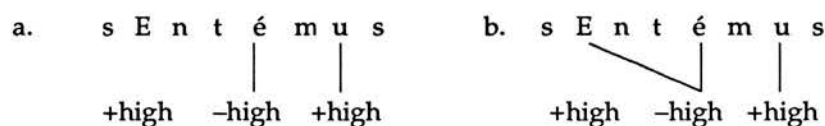
stem	áis ‘2pl pr sub’ ámus ‘1pl pr sub’	émus ‘1pl pr ind’ erémus ‘1pl fut’	ís ‘2pl pr ind’	gloss
sEnt-	sintáis	sentémus	sintí:s	“feel”
bEb-	bebámus	bebémus	bibí:s	“drink”
kOx-	koxámus	koxerémus	kuxí:s	“take”
sal(g)-	salgáis	salémus	salí:s	“leave”

⁵ Gussmann’s arguments are based on Polish data. However Slowiaczek & Dinnsen (1985) provide experimental evidence which shows that the underlying contrasts which are said to be neutralized are in fact phonetically preserved.

In the first column, the stressed vowel is *a*, and no harmony occurs; the root vowels are evidently in their ‘underlying’ forms. In the second column the stressed vowel is [–high], and the root vowels agree on this specification. In the third column the stressed vowel is [+high], and again, the root vowels agree (except for the transparent vowel *a*). Kornai (1989) has claimed that this data “shows that feature changing rules are positively required in autosegmental phonology”. The reasoning is as follows: underlying high vowels can be changed to mid and underlying mid vowels can be changed to high. If non-low vowels are unspecified for height, acquiring it from the application of a spreading rule, then there is no account of the height contrast before low vowels.

A monostratal solution involves the use of morphologically conditioned (or lexically determined) defaults. Instead of specifying an underlying value for vowel height, this is left unspecified, and each word is assigned to one of two classes which determine their default behaviour. Casting this in the notation of autosegmental phonology, we posit a floating [+high] specification for *sEnt-* and a floating [–high] specification for *bEb-* and *kOx-*. A rule (or constraint) ensures that the floating specification contributed by the stressed vowel is associated to all other non-low vowels. (This proposal has been worked out in greater detail by Vago 1988). Consider the case for the word *sentémus* shown in (1–12). The structure on the left is the ‘underlying’ form and that on the right is the output of the hypothesized spreading process.

(1–12)



This autosegment is only linked to the root if no other specification is given. Note that (1–13b) violates part of the well-formedness condition (1–3a), because it has the unassociated autosegment [+high]. As we shall see later, the temporal extent of the gesture corresponding to an autosegment is only constrained by the temporal properties of the prosodic structure to which the autosegment is linked. It is reasonable to assume that such temporally unconstrained objects have no determinate realization. This situation enables us to reconstruct a kind of deletion known as STRAY ERASURE

A similar argument has been proposed by Poser (1982) for the existence of feature-changing harmony. His analysis concerns sibilant harmony in Chumash where, in general, the sibilants of a word are either all *s* or all *š*. Consider the following examples (Poser 1982:132–3).

- The rightmost sibilant determines the palatality of all other sibilants in the word.
Poser argues:

As with the analysis of Spanish harmony, it is possible to effect morphologically conditioned defaults using floating autosegments (see Avery & Rice 1989:144 for a similar treatment). Each sibilant will have a floating autosegment for the harmonizing feature (say *anterior*), and non-sibilants will not carry a specification for the feature. (This assumes that *anterior* is only used to distinguish sibilants.) The data is now accounted for by the observation that all sibilants are associated to the rightmost *anterior* feature. For example, consider the word *kšunotš*. Its underlying representation appears in (1–14a) and its surface representation in (1–14b).

- We leave the discussion of objections to the monostratal approach here. We have

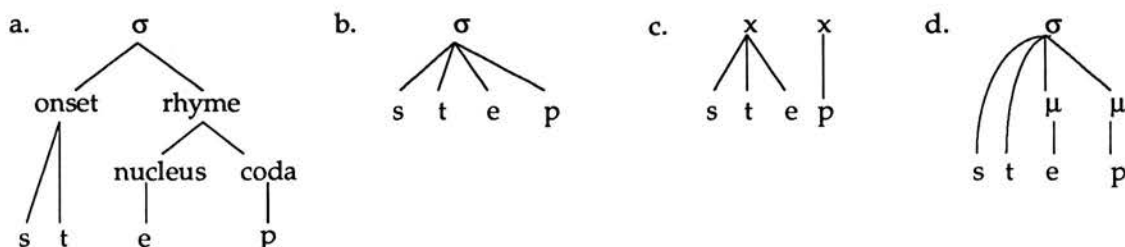
seen examples of neutralization, rule ordering and feature changing⁶. Although much more could be said, it should be clear that these objections are sometimes flawed and sometimes groundless, and that the debate is not as settled as some have thought.

Prosodic structure adds another dimension to the study of explanation in phonology—it is the subject of the next section.

1.3 Prosodic Phonology

No consensus has yet emerged about the precise role played by the syllable in phonology. For some (e.g. Firth) it was a fundamental unit, others (e.g. Chomsky & Halle 1968) have simply employed a special symbol to indicate syllable boundaries. Amongst those who gave the syllable full status there remains debate about the details of its internal structure. For example, the following structures (here exemplified for the word *step*) have been defended (e.g. Fudge 1969 (a), Kahn 1976 (b), Hyman 1984 (c), Hayes 1989 (d)).

(1–15)



The major claim of the onset-rhyme model in (1–15a), which dates back ultimately to Pike & Pike (1947) is the existence of a rhyme constituent. This claim is defended vigorously by Fudge (1987)⁷ in the face of Clements & Keyser's (1983) proposal, itself an extension of Kahn's. Hyman's model, on the other hand, has no rhyme constituent, but recognizes moras as fundamental (as does Hayes' model). A major role of the mora is in the representation of syllable weight, whereby light syllables

⁶Deletion has not been addressed directly here because of its trivial reconstruction as 'alternation with zero', following Hudson (1980). A rule of the form "delete x in the context ϕ " can be replaced with the generalization: " x appears as its zero allophone in the context ϕ ".

⁷ This is discussed further in chapter 3 section 3.

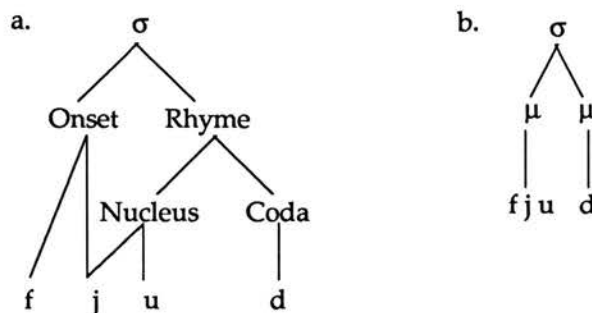
contain one mora and heavy syllables contain two. The familiar open/closed distinction is orthogonal to the light/heavy distinction, as (1–16) illustrates for *ta*, *tat* (light version), *taa* and *tat* (heavy version). The notation adopted in (1–16) and used from now on combines that of Hyman and Hayes, for reasons to be given below.

(1–16) (a) light, open (b) light, closed (c) heavy, open (d) heavy, closed



The strongest arguments in favour of the rhyme are based on the observation that more phonotactic constraints exist between the syllable peak (or nucleus) and the coda than exist between the peak and the onset (at least in English, Fudge 1987). However, it is not necessary to posit a rhyme constituent in order to express these constraints. After all, the syntactic constraint that English subjects and verbs must agree in person and number does not require the left-branching structure $[[NP\ V]\ NP]$. Although other arguments have been advanced in favour of a rhyme constituent, the evidence is not completely one-sided. Arguments have been advanced for grouping the onset and nucleus into one unit, separate from the coda. For example, Goldsmith (1990:125) postulates that “there is a maximum of one appearance of each distinctive feature over the onset-nucleus span”. Further evidence can be found in the ambiguous nature of glides. If two constituents are posited then additional complexity is the consequence (e.g. Anderson 1988 suggests that glides are initially attached to the nucleus but derivationally associated with the onset). The options are depicted in (1–17).

(1-17)

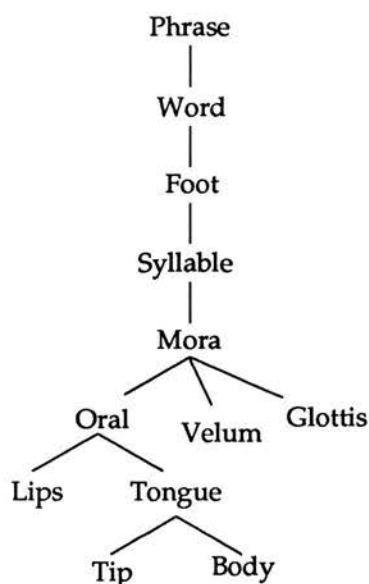


Recent phonetic evidence also supports the division of the syllable into an onset-nucleus constituent and a separate coda constituent:

Initial consonant clusters not only align their centres to the vowel, but in addition are almost completely overlapped by the vocalic gesture (which starts at the achievement of target of the first consonant). Final consonants, however, are aligned so that the target of the consonant is first attained just as the vocalic gesture is turned off. Thus, the target portions of final consonants are produced in their own time frame, while the target portion of initial consonants overlap the time frame for the vowel. (Browman & Goldstein 1988:152)

Because of the present interest in temporal structure, the classification of the syllable into an onset-nucleus constituent followed by a coda constituent will be adopted here. These constituents will be called the onset mora and the coda mora respectively, where each can dominate consonants *and* vowels. (Note that Hayes' model assumes that consonants are not linked to the first mora, only the second, but no formal account of this constraint is given.)

The organization of syllables is just a small part of the highly articulated prosodic hierarchy depicted in (1-18). This is effectively a 'side-on' view. Phrases consist of a sequence of words, which in turn consist of a sequence of feet, and so on. Moras consist of a sequence of oral, velum and glottis specifications, and so on. Refinements for individual languages (e.g. see chapter 3) constrain the number of constituents of each node in this hierarchy.



This schema omits a level which is often called the CV tier, the timing tier or the skeletal tier. This deliberate omission stems from the belief that segments and segment-sized timing units are theoretically dubious constructs. This belief is shared by McCarthy and Prince (1989) who claim that “independent motivation for the segment-sized units of CV skeleton theory is difficult to come by and often, if not always, subject to plausible reanalysis. Unambiguous evidence for segment-sized skeletal units is nonexistent”. Further arguments against the segment as the organizational unit of phonology, along with a thorough-going non-segmental model, have been provided by Griffin (1976, 1985). In the present model, the mora level is seen as a more appropriate level for stating roughly ‘segment sized’ generalizations about timing, as will be seen later.

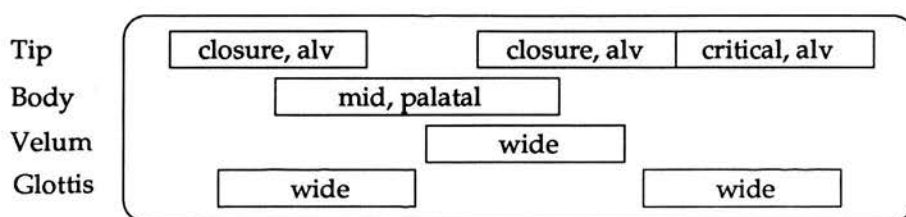
The sub-mora structure recapitulates the physiological structure of the vocal tract. The justification for this move will be evaluated extensively in chapter 3, section 1. At this stage I shall simply reproduce a passage from Sagey’s thesis which neatly encapsulates the position I am adopting. (Note that the ‘place of articulation’ feature grouping she discusses corresponds closely to the oral node in (1-18)).

Greater understanding of phonology, and a more explanatory phonological theory, result from investigating phonology hand in hand with phonetics. In phonetics are often found explanations of why phonology is the way it is. For example ‘place of articulation’ is a basic, and long-recognized, parameter in phonology. Features dealing with place of articulation form a natural class of features. Is it an accident that those features we refer

to as place of articulations features form a class in phonology? Could human language just as easily have grouped the features [constricted glottis], [coronal], and [low] into some parameter? This would be expected if the grouping into place features were purely formal, and not grounded in some way in the physical mechanism of speech. However, the grouping of features into a place constituent is not an accident, but is due to the physical mechanism of speech. ... Thus phonetics can explain why there is a unit 'place of articulation' in phonology. (Sagey 1986:17–18).

The leaves of the (inverted) tree structure in (1–18) are labelled with the active articulators: Lips, Tip, Body, Velum and Glottis. A convenient representation of utterances is the GESTURAL SCORE, which represents the temporal organization of articulatory gestures, hiding their hierarchical (phonological) organization. Here is a gestural score for the word *tense*, using Browman & Goldstein's (1989) notation.

(1–19)



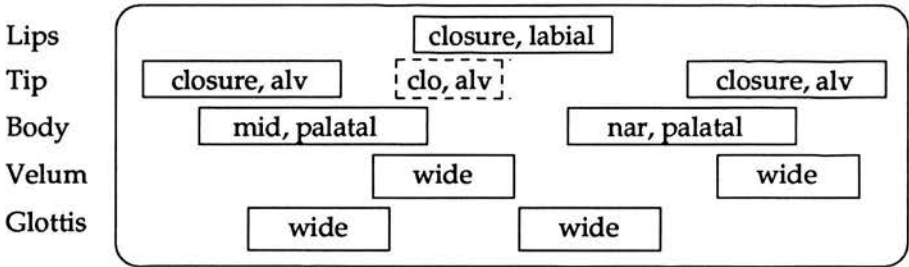
Note that the absence of articulations is not specified in this diagram (cf. Sagey 1986:65). The initial consonant *t* is formed as a result of the temporal coincidence of the first alveolar closure and the first glottal widening. The widening persists after the end of the closure period, producing a period of aspiration until the onset of voicing, by which time the tongue body is in the mid, palatal position required for the vowel. During this vowel the velum is lowered so that air can escape through the nasal cavity (heard as nasalization of the vowel). Soon there is alveolar closure again, but this time it is coordinated with the velic gesture (producing the *n*). As the velum is raised and the glottis widens, pressure builds up behind the tongue tip, which when released is heard as *t* (the so-called intrusive stop). By this time the tongue tip articulation has a critical constriction degree appropriate for *s*.

In a constraint-based setting, we can consider various utterances of *tense* as equally basic, rather than deriving 'complex' forms (e.g. [t^hɛn^ts]) from 'simpler' ones (e.g. [tɛns]) as Clements (1987) does for the intrusive stop. Language-particular phonotactic constraints can be added to eliminate some of these options. For example, the intrusive stops do not appear in South African English (Fourakis & Port 1986).

A constraint could ensure that the glottal widening of *s* does not commence before the alveolar closure commences. Further constraints can be added to account for individual differences in coarticulation (Nolan 1985).

In a similar vein, (1–20) shows the nasal assimilation which occurs in utterances of *ten pin* ([tempin]), where the tongue tip closure (indicated by a dashed box) is hidden acoustically⁸. The alveolar closure still occurs, but its acoustic effects are negligible. (See Browman & Goldstein (1989:215ff) for further discussion of acoustic hiding.)

(1–20)



In this section we have seen examples of hierarchical and temporal structure. These two notions will be themes which recur frequently through the subsequent chapters. Making a clear distinction between the two will enable important generalizations to be expressed (cf. the immediate dominance / linear precedence distinction of Gazdar et al. 1985). Phonological structures will be concerned almost solely with hierarchical structure. Utterances, on the other hand, have a rich temporal structure and no hierarchical structure. The task of phonology is to relate the two.

1.4 Sign-Based Phonology

The phonological structures described above are for morphologically simple and morphologically complex forms alike. However, the two are related: complex forms are constructed out of simple forms. Moreover, this relationship may be cast in hierarchical terms, as a phrase structure tree, for example. This raises an important question, namely, how does the phonological hierarchy of the previous section relate

⁸ It is worth noting, in this connection, that a traditional analysis which employs an assimilation rule to delete place-of-articulation specification is inadequate. Assimilations are often only partial (e.g. Nolan 1986) and are closely related to the structure of the vocal tract. For example, homorganic nasal assimilation to labio-velars stops produces a velar nasal as opposed to a labial nasal (Ryder 1987, but cf. Sagey 1986:37).

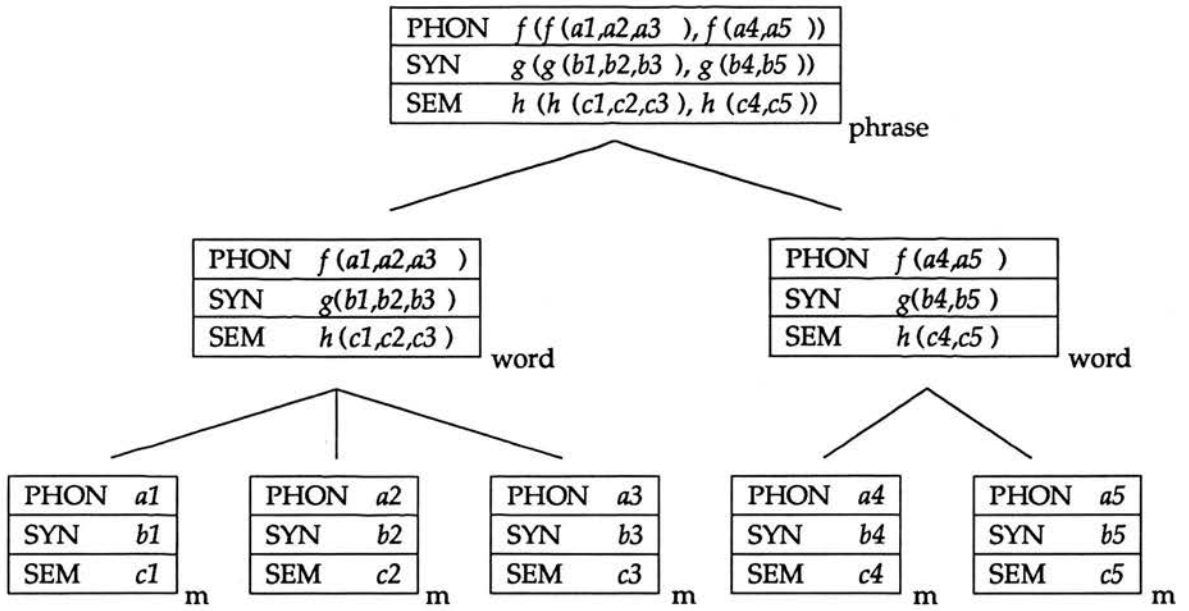
to these hierarchical morphosyntactic and syntactic structures? The traditional view is that there is just the one hierarchy, with phonological units (such as distinctive features) at the bottom, morphemes and words in the middle reaches, and phrases and sentences towards the top. The associated processing model is serial: recognition involves a phonological component, which passes its output to the morphological or syntactic component, and so on. Hockett describes this view and proposes a radical alternative:

There is a traditional view which sees phonologic and grammatic units as differing primarily as to size-level, so that the whole design of language involves but a single hierarchy: a morpheme consists of one or more phonemes; a word consists of one or more morphemes; a phrase of one or more words; and so on. The present view is radically different. Morphemes are not *composed* of phonemes at all. Morphemes are indivisible units. A given morpheme is *represented by* a certain more or less compact arrangement of phonologic material, or, indeed, sometimes by one such arrangement and sometimes by another. If we call any such representation a *morph*, then it becomes correct to say that a morph has a phonologic structure—that it consists of an arrangement of phonemes. (Hockett 1955:15)

However, this parallel view of grammatical organization was already evident in the work of de Saussure, and his notion of a linguistic sign (see Pollard & Sag 1987 for a review). In its most general sense, a sign is a pairing between a form (such as an utterance) and a meaning. Regardless of the ontological status of forms and meanings, their pairings are conventional and constitute the domain of linguistic inquiry. More specifically, I shall view a sign as having inherent phonology and semantics attributes, and that their distribution is constrained by a further attribute, called the syntax attribute⁹. This view of grammatical organization is schematized in (1–21).

⁹ The name of this attribute is somewhat misleading, as I assume it will be responsible for syntactic as well as morphosyntactic distributional constraints. This does not carry the assumption that these two kinds of distribution are being equated, however.

(1-21)



This tree represents the analysis of a phrase (the top level) into words (the middle level) and morphemes (the bottom level). The value of the attributes of each node at the word and phrase level is a function of the values of the corresponding attributes of the node's daughters. This kind of generalized (or multidimensional) compositionality has close links with Bach's (1976) rule-to-rule hypothesis and also with a large body of work in the extended Montague framework. It is formalized by Oehrle (1988).

In this study, a sign will be assumed to vary along at least three dimensions: (i) phonological content, (ii) distributional properties and (iii) contribution to semantics, as illustrated in (1-22) using a notation discussed in chapter 2. The SYN feature has further substructure, following Pollard & Sag (1987). The value of the SUBCAT feature will be assumed to be a *poset*, a view which subsumes the list and set views of HPSG and JPSG (Gunji 1986). Although the SEM attribute is included here, it will be omitted in later diagrams because natural language semantics does not impinge on the present study.

$$(1-22) \quad \left[\begin{array}{c} \text{PHON} \\ \text{SYN} \\ \text{SEM} \end{array} \begin{array}{c} \boxed{1} \\ \left[\begin{array}{c} \text{HEAD} \\ \text{SUBCAT} \end{array} \begin{array}{c} \boxed{2} \\ \boxed{3} \end{array} \right] \\ \boxed{4} \end{array} \right]$$

Signs may be combined with others in the construction of larger signs in accordance with certain combinatory rules. For two signs to be combined one must subcategorize for the other, as the schema in (1-23) requires. The ‘.’ in (1-23) represents a combination operation akin to those proposed for categorial grammar (e.g. Calder et al. 1988), except that it contains no implication for linear order or subcategorization. The \ is the set-theoretical relative complement operator.

$$(1-23) \left[\begin{array}{c} \text{PHON} \quad [1] \\ \text{SYN} \quad \left[\begin{array}{c} \text{HEAD} \quad [2] \\ \text{SUBCAT} \quad \{ \dots [4] \dots \} [3] \end{array} \right] \end{array} \right] \cdot [4] = \left[\begin{array}{c} \text{PHON} \quad [1] \\ \text{SYN} \quad \left[\begin{array}{c} \text{HEAD} \quad [2] \\ \text{SUBCAT} \quad [3] \setminus [4] \end{array} \right] \end{array} \right]$$

Observe that [3] is a set of signs having a member which unifies with [4]¹⁰. In other words, two signs may be combined if one is present in the other’s SYN|SUBCAT feature, and the resulting sign has this particular member removed. (This configuration makes explicit the assumption that signs which are syntactic functors are also phonological functors, and likewise for arguments.) Because the SUBCAT attribute contains entire signs, it may be used in conjunction with the other attributes to express distributional constraints. For example, consider the feature structure in (1-24) for the English determiner *an* (cf. Hoeksema 1985:23). Its SUBCAT list has a single element indexed by [1]. Therefore if the sign for the determiner is to combine with another sign, the latter must unify with [1].

$$(1-24) \left[\begin{array}{c} \text{PHON} \quad \langle \text{a n} \rangle + [1] \\ \text{SYN} \quad \left[\begin{array}{c} \text{HEAD} \quad \text{np[sg]} \\ \text{SUBCAT} \quad \left\langle \left[\begin{array}{c} \text{PHON} \quad \langle [+syl], \dots \rangle [1] \\ \text{SYN|HEAD} \quad \text{n[sg]} \end{array} \right] \right\rangle \end{array} \right] \end{array} \right]$$

This sign expresses the constraints that (i) whatever follows *an* must begin with a vowel and (ii) apart from the determiner, a noun phrase requires a noun.

This view of grammatical organization has consequences for the lexicon. The lexicon is essentially a store of morphemes, where each morpheme is a sign. Morpheme structure constraints can then be viewed as generalizations over the lexicon (as will

¹⁰Here I assume that [4] has an empty SUBCAT set.

be discussed in chapter 3 section 5). Compare this with the traditional view of the lexicon dating back to Bloomfield: “the lexicon is really an appendix of the grammar, a list of basic irregularities” (1933:274, cited by Hoeksema 1985:2, which Hoeksema aptly sums up as the ‘junk yard’ view.) Lexical generalizations apply to signs which happen to be morphemes; other generalizations will apply to words and phrases. The details of this organization are not important to the present study, and I shall assume my proposals will be compatible with the proposals of Karttunen (1986), Pollard & Sag (1987:191–218), Gibbon (1990) and Cooper (*forthcoming*) in this area¹¹.

1.5 Overview of Thesis

This chapter has seen the elaboration of four key aspects of the work reported in this thesis. The first section pointed out formal shortcomings of some current work in phonology and outlined an alternative model-theoretic framework. This will be discussed in detail in chapter 2. The second section presented the monostratal approach to linguistic description and responded to some of the criticisms which have appeared. The third section discussed a number of proposals concerning the hierarchical and temporal structure of utterances, focusing on the moraic model of syllable structure and Browman & Goldstein’s (1989) articulatory phonology. Chapter 3 builds on this discussion, showing how observations about hierarchical and temporal structure can be expressed in the formalism of chapter 2. The fourth section presented the sign-based approach and showed how it grew out of a diverse range of thinking. There is a fifth source of motivation (not discussed above) which pervades this study, namely the benefits which arise once a theory is computationally interpretable. Some aspects of a computational framework are presented in chapter 4, although this needs to be worked out more extensively.

The cornerstone of this whole endeavour is the description/object distinction, also known as the syntax/semantics distinction¹² or the type/token distinction (e.g. Brom-

¹¹ There are a number of points of contact between this view of lexical structure and the ‘deductive databases’ of computer science (Minker 1987), where lexical rules correspond to clauses and lexical generalizations to integrity constraints.

¹² Note that it is the syntax and semantics of the metalanguage, and not the object language which is being referred to here.

berger 1988). Phonological representations are *descriptions* of utterances. Along with this distinction comes a conflation of the representation/rule distinction, which has been widely assumed to be fundamentally important in phonology (witness Anderson 1985, criticized along these lines by Broe 1988).

Chapter 2

A Logical Foundation for Phonology

The phonological literature is rich with conventions for representing linguistic structures in graphical form. As already discussed in chapter 1, the definition of such notational conventions has two parts. First, it is necessary to be able to say of any diagram whether it is well-formed (with respect to the convention) by providing a syntax. Second, since notations are essentially *about* something, it is necessary to be explicit about their interpretation by providing a semantics. However, these conventions are usually defined by giving a small number of examples (i.e. an informal syntax) and a prose description of their interpretation (i.e. an informal semantics). The careful reader of such definitions may justifiably be concerned about things being ‘slipped under the table’, as it were.

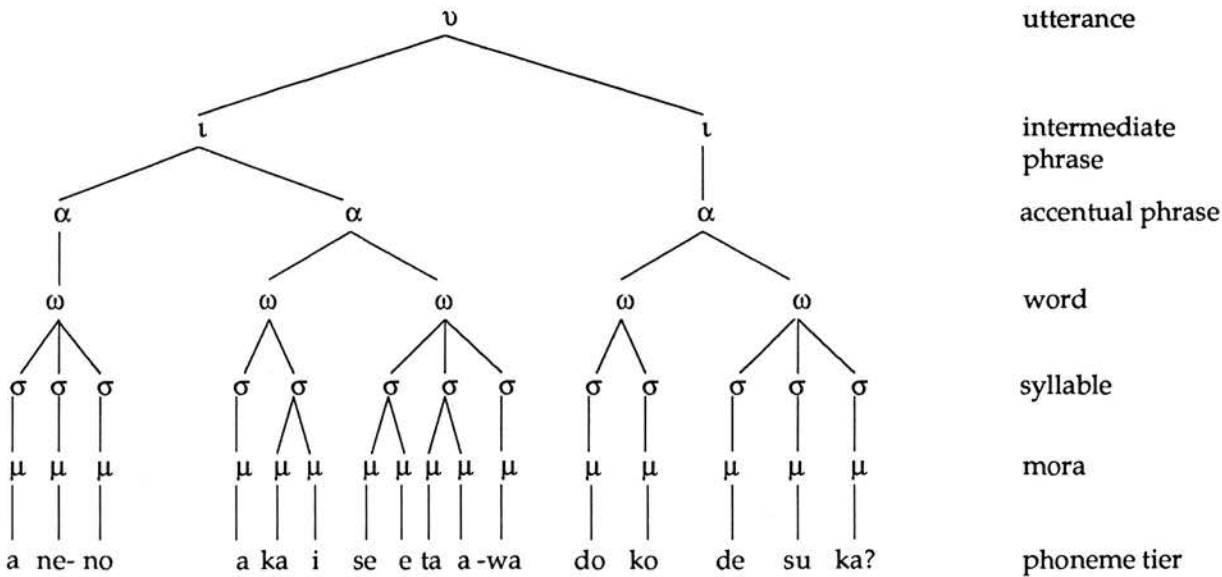
The aim of this chapter is to present a simple logical framework in which notational proposals can be expressed, evaluated and applied. Crucially, it will be possible to provide notational devices with a formal syntax and a formal semantics in this setting. Phonological representations depict various linguistic ‘objects’, such as syllables, moras, place nodes and so on, along with various temporal and structural relations between them. The first section concerns the objects themselves. The second and third sections discuss the hierarchical and temporal relations between these

objects. The fourth section explores the interaction of temporal and hierarchical relations. A second graphical notation, the feature matrix, is described in the fifth section and it is shown how descriptions of trees also describe feature matrices. After some abbreviatory devices are defined (section six) phonological rules are discussed (section seven). Throughout these seven sections a phonological description language \mathcal{L} will gradually be built up. In section eight the various parts are put together in a classical first-order theory. The ninth section shows how phonological descriptions may be combined, and section ten concludes the chapter.

2.1 Sorts

Consider the prosodic structure diagram in (2-1). It consists of a collection of labelled nodes and a collection of lines between those nodes. For now we shall concentrate on the nodes and leave the lines until later. Notice that the labels are not unique. For example, there are thirteen σ nodes, expressing the fact that the utterance contains thirteen syllables.

(2-1) **Pierrehumbert & Beckman's (1988:118) prosodic structure for a Japanese utterance:**



(Lines between syllables and consonants have been omitted from this diagram.) Our description language \mathcal{L} has variables ranging over the nodes of graphs like (2-1). These variables are represented using the symbols x , y and z (possibly subscripted). We can restrict the range of a variable by giving it a SORT. For example, we shall say x is a syllable if and only if (iff) $\text{syl}(x)$ is true. Similarly for moras: the predicate $\text{mora}(x)$ is true iff x is a mora¹. We can do this for each level of the hierarchy in (2-1).

Notice also that each node of (2-1) has only one label. It would not be possible for a node to be simultaneously labelled σ and μ since syllables and moras are distinct entities. In other words, the expression $\text{syl}(x) \wedge \text{mora}(x)$ is always false, regardless of which node x denotes. Another way of writing this is given in (2-2a). A logically equivalent formulation appears in (2-2b).

- (2-2) a. $\forall x \text{ syl}(x) \rightarrow \neg \text{mora}(x)$
 b. $\forall x \text{ mora}(x) \rightarrow \neg \text{syl}(x)$

The lowest level of the above prosodic structure is the 'phoneme tier'. Each element of this tier is a phoneme. However, unlike the other levels of prosodic structure, it is not the case that all nodes on this level have the same label. For example, there are seven 'a' labels and four 'e' labels. Continuing with the sort predicates, $a(x)$ could pick out just those nodes which are labelled with an 'a', and $e(x)$ could pick out those nodes labelled 'e', and so on.

Here, for the first time, is a situation where it makes sense to think of nodes as having more than one label. In addition to the existing labels, all nodes in the phoneme tier might be labelled π , for example, to indicate that they are phonemes. A predicate $\text{phoneme}(x)$ could pick out just those nodes in the phoneme tier. The situation can be expressed using the following (partial) formula:

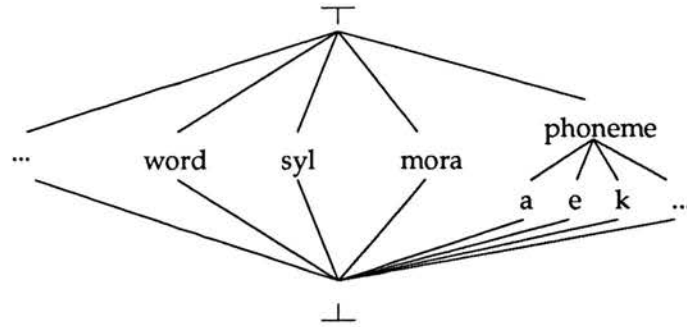
- (2-3) $\forall x \text{ phoneme}(x) \equiv a(x) \vee e(x) \vee k(x) \vee \dots$

This formula states that something is a phoneme if and only if it is also an a or an e or a k or a ... (and so on). These relationships between sorts may be expressed

¹ This sort system concerns the nodes of a tree. It is not intended to give us a way of saying, for example, that the English word 'a' is simultaneously a phoneme, mora and a syllable.

graphically using a LATTICE².

(2-4) A Partial Sort Lattice for Prosodic Structures:



The top-most node in the lattice is labelled \top . By convention, \top is the most general description and in the present context it describes all prosodic structure nodes. The downwards branching arcs can be thought of as expressing disjunctions. For example, a node described by \top (i.e. all nodes) can be also described as a phoneme or a mora or a syllable or ... (and so on). In turn, a phoneme can also be described as an *a* or an *e* or ... (and so on). The bottom-most element in the lattice is labelled \perp . By convention, \perp represents inconsistent information; it describes nothing. The lattice expresses the constraint that if a node is both a syllable and a mora, then it is also described by \perp . In other words, no node can be both a syllable and a mora. Similarly, no node can be both an *a* and an *e*, and no node can be both a *word* and a *k*.

Each line in the above lattice corresponds to an implication. For example, the line between the *a* and phoneme nodes corresponds to the following implication.

$$(2-5) \quad \forall x \ a(x) \rightarrow \text{phoneme}(x)$$

More specifically, the above lattice diagram corresponds to the following formula.

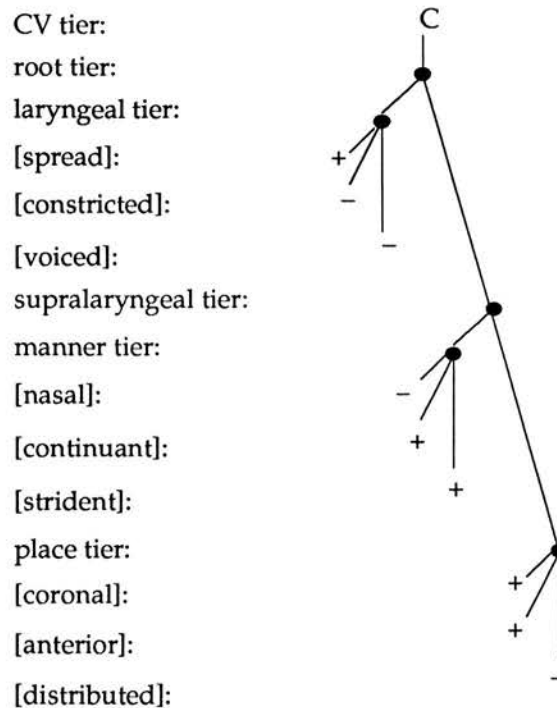
(Here we assume that the above lattice diagram is complete, ignoring the ellipses).

² Although we shall often depict such sort lattices, there will always be an algebraic structure in the background. We can view the implication relation, \rightarrow , as a reflexive, antisymmetric and transitive order on a set. More will be said about the algebraic view in section 8.

$$\begin{aligned} (2-6) \quad & \forall x \quad (\top(x) \equiv \text{word}(x) \vee \text{syl}(x) \vee \text{mora}(x) \vee \text{phoneme}(x)) \\ & \wedge (\text{word}(x) \rightarrow \neg \text{syl}(x) \wedge \neg \text{mora}(x) \wedge \neg \text{phoneme}(x)) \\ & \wedge (\text{syl}(x) \rightarrow \neg \text{mora}(x) \wedge \neg \text{phoneme}(x)) \\ & \wedge (\text{mora}(x) \rightarrow \neg \text{phoneme}(x)) \\ & \wedge (\text{phoneme}(x) \equiv \text{a}(x) \vee \text{e}(x) \vee \text{k}(x)) \\ & \wedge (\text{a}(x) \rightarrow \neg \text{e}(x) \wedge \neg \text{k}(x)) \\ & \wedge (\text{e}(x) \rightarrow \neg \text{k}(x)) \end{aligned}$$

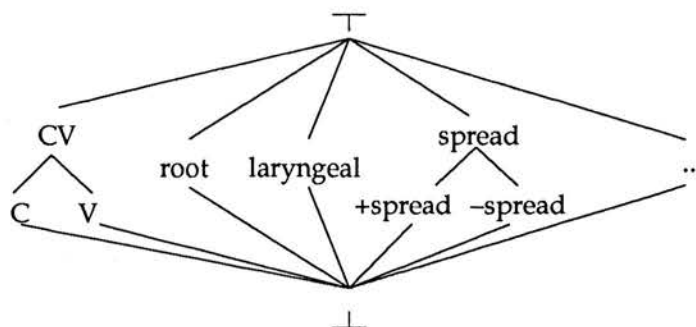
As a second example, consider the kinds of diagrams which have been proposed for representing hierarchical sub-segmental structures. An example of such a structure is given in (2-7).

(2-7) Clements' (1985:248) partial representation for [s]:



In this diagram only some of the nodes are labelled. However, we can think of the column of words to the left of the tree as being a collection of node labels. The highest unlabelled node will be considered to have the label 'root', the next 'laryngeal', the next 'spread' and so on (Sagey (1986), Hayes (1990) and others label all nodes in this way.) The sort lattice for these nodes is given in (2-8).

(2-8) A Partial Sort Lattice for Sub-Segmental Structures:



The CV sort picks out all CV-tier elements, the C sort picks out the consonants (a subset of the set picked out by CV) and so on. Note that *spread* picks out the objects which can have a specification for the [spread] feature. The sorts *+spread* and *-spread* pick out (mutually exclusive) subsets of the set picked out by *spread*. So, we can see that the purpose of sorts is to pick out various subsets of a given set of nodes³.

In this section we have seen a technique for classifying the nodes of a hierarchical structure through the use of a sort system. However, this is only a fragment of the overall picture. For example, as things stand at present, there is nothing to stop a vowel in this CV-tier from dominating sub-segmental structure containing a [-continuant] feature, which is impossible as all vowels are [+continuant]. In order to solve this problem it is necessary to postpone further discussion of sorts until after the hierarchical organization of phonological structures has been discussed.

2.2 Hierarchical Organization in Phonological Structures

In the previous section we saw two kinds of hierarchical structure, one for prosodic structure and one for sub-segmental structure. The nodes of these trees are related to each other by dominance (Clements 1985:248, Pierrehumbert & Beckman 1988:145-6).

Clements' definition is as follows:

³ The reason for not employing + and - as sorts directly is because we never want to refer to the class of nodes that have a + value as opposed to a - value of an arbitrary feature. For example, Clements' representation for [s] above has five nodes with a '+' label: *spread*, *continuant*, *strident*, *coronal* and *anterior*. These nodes have no more in common with each other than any other subset of nodes. This is in line with the widespread criticism of the α notation of SPE, which enabled rules like $[\alpha\text{spread}] \rightarrow [\alpha\text{nasal}]$, where α ranges over {+, -}. (It is considered one of the merits of the autosegmental approach that such operations are not possible.) Note also that the use of sorts *spread*, *+spread* and *-spread* implements an internal view of negation: $\neg +\text{spread}(x) \neq -\text{spread}(x)$.

(2–9) **Clements' (1985:248) definition of dominance:**

Given any two nodes M, N such that M lies on the path between N and the root of the tree, M is said to DOMINATE N; if no node intervenes between M and N, M IMMEDIATELY DOMINATES N.

The lines in Clements' structure for the segment [s] represent the immediate dominance relation. This relation is asymmetric: if M immediately dominates N then it is not possible for N to immediately dominate M. If the vertical orientation of these diagrams on the page were not significant then it would be necessary to employ arrows instead of lines in order to capture this directionality. Note also that immediate dominance is irreflexive: it is not possible for a node to dominate itself. Lastly, immediate dominance is intransitive: if M immediately dominates N and N immediately dominates P then it is not the case that M immediately dominates P. These three properties can be stated as follows, where δ is the immediate dominance relation.

(2–10) a. Immediate Dominance is Irreflexive:

$$\forall x \neg x \delta x$$

b. Immediate Dominance is Asymmetric:

$$\forall xy \ x \delta y \rightarrow \neg y \delta x$$

c. Immediate Dominance is Intransitive:

$$\forall xyz \ x \delta y \wedge y \delta z \rightarrow \neg x \delta z$$

The dominance relation δ^* is defined as the transitive closure of the immediate dominance relation. It is therefore irreflexive, asymmetric and transitive.

Now that sorts and the δ relation have been defined, it is reasonable to wonder if there is any significant interaction between the two. It turns out that there is. Notice that in both Pierrehumbert & Beckman's and Clements' diagrams above it is not possible for a line to connect nodes of arbitrary sorts⁴. For example, it is impossible for a line to connect an ω node to a μ node. Likewise, it is impossible for a root node to immediately dominate a manner node. For Pierrehumbert & Beckman's tree we employ the following constraints. (Note that some of the sorts are abbreviated from the labels.)

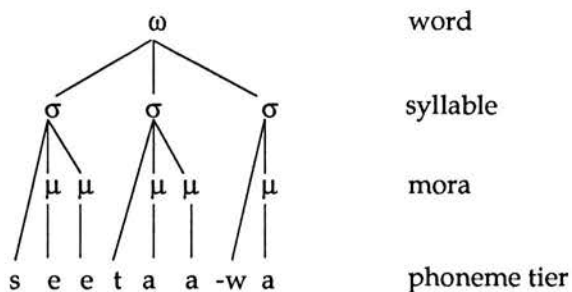
⁴ This fact is seldom made fully explicit in the phonological literature. See (Sagey 1986:33) for a rare example.

(2-11) **Appropriateness in Pierrehumbert & Beckman's tree:**

$$\begin{aligned}
 \forall xy \ x \delta \ y &\rightarrow (\text{utt}(x) \wedge \text{ip}(y) \\
 &\vee \text{ip}(x) \wedge \text{ap}(y) \\
 &\vee \text{ap}(x) \wedge \text{word}(y) \\
 &\vee \text{word}(x) \wedge \text{syl}(y) \\
 &\vee \text{syl}(x) \wedge (\text{mora}(y) \vee \text{phoneme}(y)) \\
 &\vee \text{mora}(x) \wedge \text{phoneme}(y))
 \end{aligned}$$

This constraint says of any pair of nodes x and y where x immediately dominates y that either x is an utterance and y is an intermediate phrase, or x is an intermediate phrase and y is an accentual phrase, or ... (and so on). Note that the fifth line of (2-11) is more complex than the others. If x is a syllable then y is either a mora or a phoneme. This is necessary in order to admit structures of the following form:

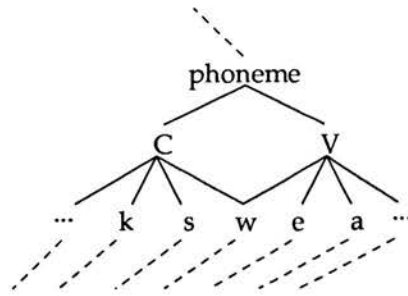
(2-12) **Pierrehumbert & Beckman's (1988:119) partial prosodic structure:**



Notice that in this structure the only kind of phonemes which are immediately dominated by a mora are vowels and the only kind immediately dominated by a syllable are consonants. The former is simply a coincidence and there are many situations where we shall want to permit moras to dominate consonants. However, the fact that syllables can immediately dominate consonants (and not vowels) is part of the definition of syllable structure in the moraic theory.

In order to express this constraint succinctly we first enrich the underlying lattice slightly to include the C and V sorts. The result is given in (2-13). Note the ambivalent status of the glide w .

(2-13) A Partial Sort Lattice for Prosodic Structures:

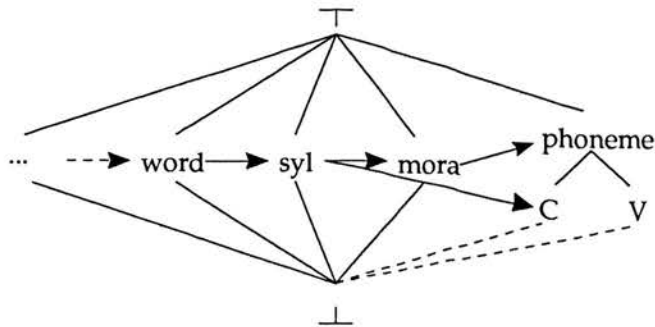


In the context of this enriched sort lattice, the constraint that the only kind of phoneme a syllable can immediately dominate is a consonant is stated in (2-14).

$$(2-14) \quad \forall xy \quad x \delta y \wedge \text{syl}(x) \wedge \text{phoneme}(y) \rightarrow c(y)$$

Thus far we have seen a range of appropriateness constraints stated as logical formulae. It is convenient to have a graphical representation of these constraints. One way to achieve this is to view δ as a binary relation on sorts. If it is possible for a node of sort s to immediately dominate a node of sort t , then we enter an arrow between s and t into the sort lattice, as shown in (2-15).

(2-15) A Graphical Depiction of Appropriateness for Pierrehumbert & Beckman's Tree:



The arrow represents the immediate dominance relation δ . The dashed arrow and dashed lines indicate elided structure. Two arrows emanate from the sort *syl*, since it can immediately dominate two different sorts of node. If a syllable immediately dominates a phoneme then that phoneme must also be a consonant.

A similar representation can be provided for Clements' tree. However, this does not go quite far enough. As has already been observed, it is necessary to prevent

a vowel on the CV tier from dominating a –continuant specification. Indeed, this is just one of many similar dependencies which cannot be expressed directly using appropriateness constraints on immediate dominance. Instead, it is necessary to employ the transitive version of this relation δ^* . The statement of the constraint appears in (2–16).

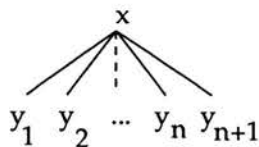
$$(2-16) \quad \forall xy \quad v(x) \wedge x \delta^* y \wedge \text{continuant}(y) \rightarrow +\text{continuant}(y)$$

We have seen that appropriateness constraints restrict the sorts of nodes which can be immediately dominated by a particular node. For example, if a root node immediately dominates another node, then that node must be either a laryngeal node or a manner node. However, this has nothing to do with the total number of nodes which a node can dominate. Given the formalism so far, the situation where a root node dominates a thousand laryngeal nodes is as acceptable as the situation where a root node dominates only one laryngeal node. Clearly we shall want to place some restrictions on the amount of branching which can occur (cf. Goldsmith 1990:18–19). The most common situation in sub-segmental structures is for there to be at most one occurrence of a given sort of node dominated by any node. This may be expressed as follows, exemplified for the laryngeal node.

$$(2-17) \quad \forall xyz \quad x \delta y \wedge x \delta z \wedge \text{laryngeal}(y) \wedge \text{laryngeal}(z) \rightarrow y = z$$

This can be generalized to an arbitrary number. Suppose that nodes of sort t can occur at most n times as constituents of a node of sort s . Then if some node dominates $n + 1$ nodes of sort s then at least two of these nodes must be identical. This may be expressed as follows.

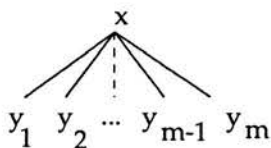
- (2-18) a. $\forall x y_1 \dots y_{n+1} \quad s(x) \wedge x \delta y_1 \wedge \dots \wedge x \delta y_{n+1} \wedge t(y_1) \wedge \dots \wedge t(y_{n+1})$
 $\rightarrow \bigvee_{0 \leq i \neq j \leq n+1} y_i = y_j$
 b.



Two y's must
be identical

Constraints of this kind impose an upper limit on branching. In certain situations it may also be desirable to be able to impose a lower limit on branching. Suppose some sort s must occur at least m times as a label on the constituents of a node. This is expressed as follows:

- (2-19) a. $\forall x \quad s(x) \rightarrow (\exists y_1 \dots y_m \quad x \delta y_1 \wedge \dots \wedge x \delta y_m \wedge t(y_1) \wedge \dots \wedge t(y_m))$
 $\wedge \bigwedge_{0 \leq i \neq j \leq m} y_i \neq y_j$
 b.



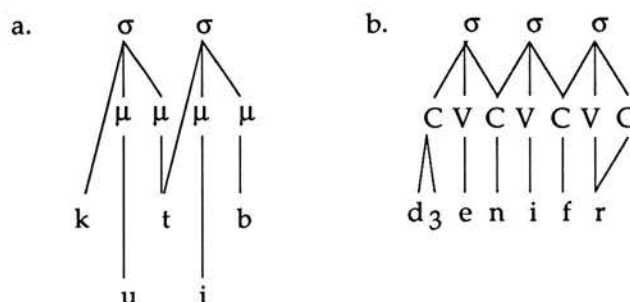
No y's are
identical

Although the constraints in (2-18) and (2-19) are rather unwieldy, the value of n typically will not exceed 3 and will normally be 1, and the value of m typically will not exceed 1. (An abbreviatory notation for these constraints will be given in section 6. Exemplification can be found in chapter 3 section 4.)

Now that constraints on the form and number of branches in these hierarchical structures have been examined, the question remains as to what exactly these structures are. So far we have seen prosodic and sub-segmental structure represented using trees. However, it is widely recognized that phonological structures are not trees, as the following examples illustrate (McCarthy & Prince 1989, Clements & Keyser

1983).

(2-20)



In both structures there are nodes which are dominated by more than one node. For example, the *t* in (2-20a) and the *r* in (2-20b) are both dominated by two nodes. Structures containing nodes like these will be called re-entrant. Therefore no constraint which prevents the possibility of re-entrancy will be built into the system. If it is required for certain grammars then such a constraint may easily be added. The constraint in (2-21) states that if two nodes *x* and *y* dominate a given node *z* then those two nodes cannot be distinct (cf. Pierrehumbert & Beckman 1988:153):

(2-21) **A constraint to prevent re-entrancy:**

$$\forall xyz \ x \ \delta \ z \wedge y \ \delta \ z \rightarrow x = y$$

This constraint may be required for certain grammars; it is not proposed as a universal constraint. It may even be that we wish to state the above constraint for particular kinds of nodes only, and on a language-specific basis. For example, consider the parameters proposed by Paradis & Prunet (1989:323) for the Fula language:

(2-22) Parameters: nodes allowed to spread:

- i. articulator nodes? Fula: *no*
- ii. Place nodes? Fula: *yes*

It is only necessary to add a constraint for (2-22i), as the default situation is for spreading to be allowed. A constraint which prevents articulator nodes from spreading is given in (2-23).

(2-23) $\forall xyz \ x \ \delta \ z \wedge y \ \delta \ z \wedge \text{articulator}(z) \rightarrow x = y$

The above discussion concerned the re-entrancy property. There are other properties of trees which we may or may not require. For example, trees must be rooted. A root

is a special node from which all others are accessible by traversing arcs (which are normally assumed to be directed). Trees are also connected, a property which follows from the fact that trees are rooted. Phonologists regularly employ non-rooted, non-connected structures when dealing with a phenomena known as extrametricality or extrasyllabicity. Therefore we do not wish to stipulate that phonological structures must be rooted or connected⁵.

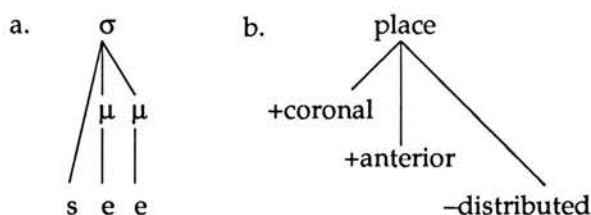
A final property is acyclicity. This property is assumed by every theory of phonological structure I am aware of. Fortunately, it follows directly from the irreflexivity and transitivity of the dominance relation δ^* .

In this section the significance of the vertical orientation of phonological diagrams has been dealt with. Two dominance relations δ and δ^* have been introduced, and it has been shown how these may be used to represent hierarchical structure. Constraints on the interaction of δ and δ^* with the sort system have been used to express a number of observations about the appropriateness of certain specifications in certain structural positions. A treatment of vertical orientation is only half of the problem. Now we turn to some issues concerning the interpretation of horizontal orientation.

2.3 Temporal Organization in Phonological Structures

Let us examine again the two running examples. A small subpart of each is reproduced below. (2-24a) depicts a syllable containing an onset and two moras. (2-24b) depicts the place subtree for the segment [s].

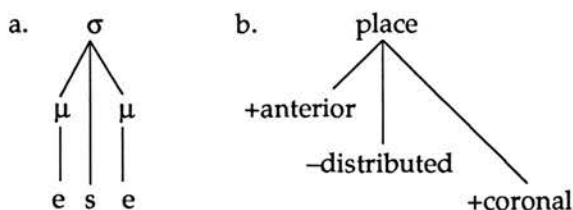
(2-24)



⁵ Such a property may be required at certain stages of a derivation; an example is discussed at the end of chapter 3 section 4.

Now consider the same trees with the horizontal arrangement permuted. (2-25a) is nonsensical because the onset is in the middle of the syllable, an impossible situation. However, (2-25b) is perfectly well-formed. Moreover, (2-25b) describes the same class of linguistic objects as (2-24b).

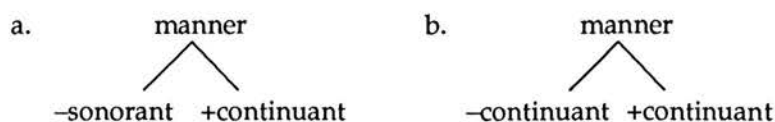
(2-25)



Therefore, the interpretation of (2-24a) and (2-25a) is different from the interpretation of (2-24b) and (2-25b). Clements (1985:226) calls the former kind of interpretation *sequential* and the second kind *simultaneous*. In the two (a) examples, left-to-right ordering on the page corresponds to temporal ordering. The onset is left of the moras because the onset is the first part of a syllable to be uttered. In the (b) examples left-to-right ordering is insignificant. An [s] segment is only created when the three daughters of the place subtree are 'articulated' simultaneously.

This difference of interpretation lies in the identity of the nodes themselves. If distinct nodes have the same label (which may be implicit in the case of Pierrehumbert & Beckman's phonemes) then one must precede the other. In a prosodic tree it tends to be the case that all daughters of a node share the same label and are linearly ordered. In a sub-segmental tree it is more usual for all daughters of a node to have distinct labels and to be 'coarticulated'. However, sub-segmental structures do not exclusively involve coarticulation. For example, (2-26a) is a (partial) manner subtree for [s] and (2-26b) is a (partial) manner subtree for an affricate (Sagey 1986:28, Hayes 1990:60).

(2-26)



In (2–26a) the daughters are coarticulated and left-to-right ordering is insignificant. However, in (2–26b) the daughters are not coarticulated; the part of an utterance corresponding the left daughter precedes that corresponding to the right daughter. Again, this difference of interpretation is accompanied by a difference in the patterning of sorts. In (2–26a) neither daughter shares a common sort, whereas in (2–26b) both daughters have the sort *continuant*. This view is compatible with Sagey’s discussion; she states that “it is only branchings to specifications on a single tier that are phonologically ordered” (Sagey 1986:28).

One way to signal this difference formally is to introduce two new binary relations in addition to the δ relation. These relations are precedence, \prec and overlap, \circ ⁶. Van Benthem (1983) provided a detailed study of the properties of these temporal relations, and their application to phonology has already been discussed at length by Bird & Klein (1990). The relevant section of the latter is included in Appendix A, and so the following properties of these relations will be assumed without further justification.

- (2–27) a. Overlap is reflexive:
 $\forall x \ x \circ x$
 b. Overlap is symmetric:
 $\forall xy \ x \circ y \rightarrow y \circ x$
 c. Precedence is asymmetric:
 $\forall xy \ x \prec y \rightarrow \neg y \prec x$
 d. Precedence is disjoint from overlap:
 $\forall xy \ x \prec y \rightarrow \neg x \circ y$
 e. Precedence is transitive (through overlap):
 $\forall wxyz \ w \prec x \wedge x \circ y \wedge y \prec z \rightarrow w \prec z$
 f. Time is linear:
 $\forall xy \ x \prec y \vee x \circ y \vee x \succ y$

One of the important consequences of these axioms is known as the NO CROSSING CONSTRAINT—already discussed in chapter 1—which requires that the lines of diagrams do not cross.

(2–28) **The No Crossing Constraint:**

$$\neg \exists wxyz \ w \prec x \wedge y \prec z \wedge w \circ z \wedge x \circ y$$

⁶ These relations are temporal orderings on nodes, where nodes are to be thought of as temporal events (in the sense of Bird & Klein 1990; see Appendix A). As we shall see later, nodes and temporal events likewise are the denotations of the terms of the language \mathcal{L} .

Although the temporal relations \prec and \circ can hold between any pair of nodes, they will generally be stated explicitly just for nodes that are sisters. These temporal relations will be used between sister constituents of a node. For (2–26), suppose a is the node at the end of the left branch and b is the node at the end of the right branch. For (2–26a) we have the constraint $a \circ b$ while for (2–26b) we have the constraint $a \prec b$.

Those constituents which have the same sort will be required to stand in a linear ordering. This is formulated in (2–29). Note that (2–29) is not an axiom but an axiom schema, which stands in place of a number of axioms. In fact, there will be an instance of (2–29) for each sort s .

(2–29) **Linearity axiom schema:**

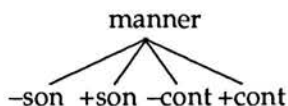
$$\forall xy \ s(x) \wedge s(y) \rightarrow x \prec y \vee x = y \vee x \succ y$$

Suppose x and y have the same sort s . Then either x precedes y or x equals y or x follows y . Consider (2–26a), where the left node (a) has the labels ‘sonorant’ (implicit) and ‘–sonorant’ (explicit), and the right node (b) has the labels ‘continuant’ and ‘+continuant’. None of the sorts of a coincide with any of the sorts of b , and so none of the axioms of (2–29) apply to (2–26a). In contrast, the two daughter nodes of (2–26b) share the sort *continuant*. Therefore they must either be equal or one must precede the other. However, if they were equal then there would be a node which is simultaneously *+continuant* and *–continuant*, which is a contradiction. Therefore they must be linearly ordered.

The linear ordering axioms have an important articulatory phonetic basis. They state that a given articulator cannot do two different things simultaneously. Here is another example of the way the model-theoretic approach works: observations about the properties of the domain are stated in the system. This view contrasts with a prevalent attitude amongst phonologists where such moves are interpreted as importing phonetics into phonology, causing undesirable duplication. As already mentioned in chapter 1, the latter view seems to be based upon a category mistake.

The linear ordering axioms say nothing about the fact that constituents having different sorts must overlap, as in the case of the +anterior, –distributed and +coronal feature specifications under the place node for [s]. In the general case, we cannot *require* constituents having different sorts to overlap. For example, consider the following partial structure.

(2–30)



The above linear ordering axioms ensure that the two continuant nodes do not overlap and that the two sonorant nodes do not overlap. If we required nodes having different sorts to overlap then we would have a situation where both continuant nodes overlap both sonorant nodes. This leads directly to a violation of the no-crossing constraint.

This putative constraint leads to another logical contradiction. Consider the two sister nodes above (say *a* and *b*) which have the sorts –*sonorant* and +*sonorant* respectively. So we have –*sonorant*(*a*) and +*sonorant*(*b*). Since they share the sort *sonorant* they must be linearly ordered. We have seen above that $a \neq b$ and so either $a \prec b$ or $a \succ b$. Now if sister nodes of different sorts were required to overlap then we would have $a \circ b$. Given the disjointness of precedence and overlap, this is a contradiction.

The question of the need for further constraints on the interaction of δ , \circ and sorts remains open. Furthermore, it is also an open question as to which further temporal relations should be adopted. For example, the following relations have been discussed in the literature (van Benthem 1983, Allen 1983, Schmeidel 1988): includes, is-conterminous-with, starts-first, ends-first, meets, and many more.

Now that we have discussed the temporal relationship between sister constituents, it is natural to wonder what temporal relationship, if any, holds between a node and its constituents⁷. At an intuitive level, it is reasonable to suppose that the temporal extent of, say, a phrase is co-extensive with the combined temporal extents of its

⁷ Recall that the nodes are to be thought of as temporal events.

constituents. The various levels of hierarchical structure are all located in one and the same flow of time. However, it is not a trivial matter to add formal content to these intuitions, as the next section shows.

2.4 The Interaction of Hierarchical and Temporal Structure

So far in this chapter we have seen the definition and application of sorts, two dominance relations (δ , δ^*) and two temporal relations (\prec , \circ) and we have seen interactions between sorts and the dominance relations, and interactions between sorts and the temporal relations. In this section the interactions between dominance and temporal relations are discussed.

None of the hierarchical structures presented above contained crossing lines. This seems to be a general requirement, even in situations where re-entrancy is employed. However, nothing about the dominance relation currently prevents such a possibility from arising. A simple, intuitive way to prevent this is to add the following constraint.

(2-31) **Locality Constraint:**

$$\forall xy \ x \ \delta^* \ y \rightarrow x \circ y$$

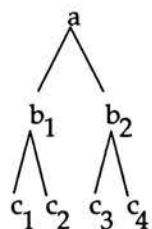
Now any nodes related by dominance overlap. A situation where the lines of dominance cross reduces immediately to a violation of the no-crossing constraint. This constraint has further motivation: the utterance of a word or phrase temporally overlaps the utterances of the constituents of the word or phrase. For example, the period of time occupied by an utterance of the word 'cat' temporally overlaps the period of time occupied by the utterance of 'c' which forms a part of the utterance of 'cat'. Hayes (1990:44) has adopted a constraint very similar to (2-31), his 'percolation convention'.

The locality constraint brings dominance and overlap into a rather close relationship. However, they are not to be conflated, as there are numerous cases of overlap between nodes which are not related by dominance, such as in most sub-segmental structures.

This constraint has important consequences for the inheritance of precedence. Con-

sider the following structure, resembling the prosodic trees we have seen already.

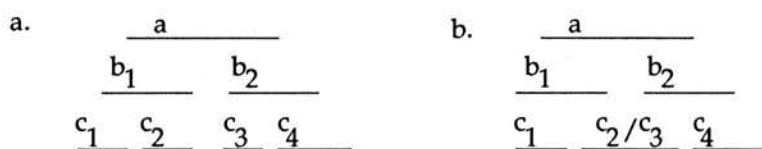
(2-32)



Each level of this tree is assumed to consist of nodes having the same sort. Suppose that b_1 and b_2 are distinct. Then the relevant linearity axiom ensures that they cannot overlap. Suppose further that $b_1 \prec b_2$. In a similar fashion, suppose that $c_1 \prec c_2$ and $c_3 \prec c_4$. What, then, is the relationship between the c nodes dominated by b_1 and the c nodes dominated by b_2 ? We know that it must be \prec , $=$ or \succ . From the no-crossing constraint, we know that none of the nodes dominated by b_2 precede any of the nodes dominated by b_1 (since both b_1 and b_2 overlap all of their constituents). Therefore, each node dominated by b_1 must either precede or equal each node dominated by b_2 .

It turns out that there are exactly two possibilities. These are indicated schematically in (2-33), where the temporal extent of the nodes is indicated explicitly.

(2-33)



The picture in (2-33a) is intended to show that both c_1 and c_2 completely precede both c_3 and c_4 . The picture in (2-33b) is ambiguous between the possibility of c_2 and c_3 overlapping, and the possibility of c_2 and c_3 being identical. The following discussion will assume the latter, without loss of generality.

Now we can return to the conclusion stated above that each node dominated by b_1 must either precede or equal each node dominated by b_2 . In the case of c_2 and c_3 both options are permitted. In other words, it is allowable for b_1 and b_2 to share a node if it is the rightmost of b_1 's daughters and the leftmost of b_2 's daughters.

Consider next the case of c_1 and c_3 . Suppose first that $c_1 = c_3$. Then $c_3 \prec c_2$. However, we know that $c_2 \circ b_1 \prec b_2$ and so one of the axioms can be applied to infer that $c_3 \prec b_2$. But $b_2 \circ c_3$ by the locality constraint, which is a contradiction. Therefore it must be the case that $c_1 \prec c_3$. By a similar argument $c_2 \prec c_4$. From the transitivity of \prec we can also infer that $c_1 \prec c_4$. The conclusion is as follows: structure sharing only occurs at the edges of subtrees. In graphical terms, this is equivalent to the requirement that the lines of dominance do not cross.

In this section we have seen areas of interaction between hierarchical and temporal structure. Hierarchical, temporal and sort structure all interact in the following situation. Suppose that the ordering within a constituent is fully predictable from the sorts. An example of this comes from syllable structure, where onsets must precede moras. In this case, we could employ the constraint in (2-34)⁸.

$$(2-34) \quad \forall xyz \quad x \delta y \wedge \text{onset}(y) \wedge x \delta z \wedge \text{mora}(z) \rightarrow y \prec z$$

This concludes the discussion of the motivation and interaction of sortal, hierarchical and temporal information. Now we turn to the representation of this information.

2.5 Another Graphical Notation: The Feature Matrix

The kinds of structures we have seen so far have involved lines between labelled nodes. However, it is sometimes convenient to be able to think of structures as consisting of labelled lines between unlabelled nodes. Of course the most general case is to have a mixture of the two: labelled nodes and labelled lines.

We shall adopt the following convention (later to be replaced by an abbreviatory device (2-47)).

$$(2-35) \quad \forall xy \quad s(x,y) \text{ iff } x \delta y \wedge s(y)$$

This states that if a *node* has a label 's' then the *lines* of dominance reaching that node from above are also labelled 's'. (These predicates may be viewed as partitions

⁸ This constraint has connections with the linear precedence statements and the 'ECPO' hypothesis of GPSG (Gazdar et al. 1985).

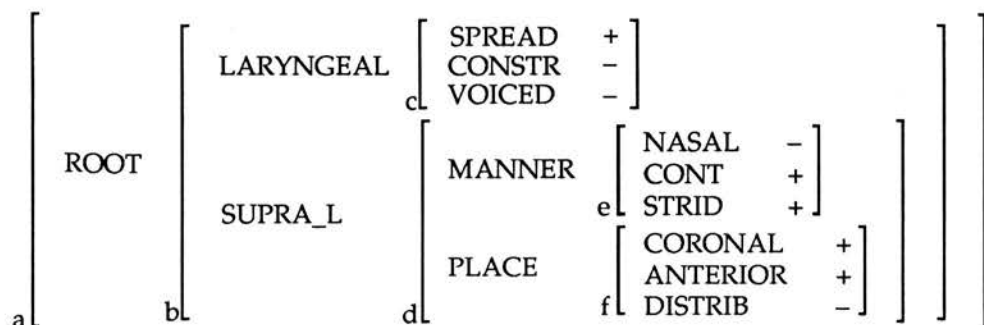
of the δ relation.) An example for part of Clements' (1985:248) representation of [s] appears below:

(2-36)



This mapping permits the trees we have seen to be depicted as feature matrices (Johnson 1988, Kasper & Rounds 1986, 1990)⁹. For example, Clements' representation for [s] can be expressed (in full) using the following feature matrix.

(2-37)



This matrix notation encodes the same information as Clements' tree diagram. Each pair of brackets and its contents is called a feature matrix. The most deeply nested entities, namely the + and - symbols, are also feature matrices, but of a special kind. They are called ATOMIC FEATURE MATRICES. The + and - symbols will be thought of here as abbreviatory, standing in place of the fuller (redundant) forms such as +spread.

A formula describing both Clements' picture and the above feature matrix is given in (2-38). The indentation is provided to aid visual comparison with (2-37).

⁹ A minor caveat is necessary here. Recall that nodes can be thought of as having more than one label. This is because a node picked out by a sort (say) +spread is also picked out by the more general sort spread. In this case, the most general label only will become the label of the incoming arc(s).

$$\begin{aligned}
(2-38) \quad & c(a) \wedge \\
& \text{root}(a,b) \wedge \\
& \text{laryngeal}(b,c) \wedge \\
& \text{spread}(c,c_1) \wedge +\text{spread}(c_1) \wedge \\
& \text{constr}(c,c_2) \wedge -\text{constr}(c_2) \wedge \\
& \text{voiced}(c,c_3) \wedge -\text{voiced}(c_3) \wedge \\
& \text{supra}_l(b,d) \wedge \\
& \text{manner}(d,e) \wedge \\
& \text{nasal}(e,e_1) \wedge -\text{nasal}(e_1) \wedge \\
& \text{cont}(e,e_2) \wedge +\text{cont}(e_2) \wedge \\
& \text{strid}(e,e_3) \wedge +\text{strid}(e_3) \wedge \\
& \text{place}(d,f) \wedge \\
& \text{coronal}(f,f_1) \wedge +\text{coronal}(f_1) \wedge \\
& \text{anterior}(f,f_2) \wedge +\text{anterior}(f_2) \wedge \\
& \text{distrib}(f,f_3) \wedge -\text{distrib}(f_3)
\end{aligned}$$

Rather than considering (2-38) and (2-37) to be notational variants, (2-37)—like the phonological diagrams given earlier—will be viewed as an object in the domain. We shall say that (2-38) *describes* (2-37). More specifically, the *denotation* of the constant a is the whole of (2-37), the denotation of the constant b is the outermost-but-one feature matrix in (2-37), and so on. Note that the constants a, b, c, d, e and f are marked in (2-37) as subscripts. The other constants in (2-38) could likewise be marked as subscripts on the individual + and – atomic feature matrices of (2-37). The relationship between a and b is $\text{root}(a,b)$, which is equivalent to stating that a dominates b and that b has the sort *root*. The denotation of the rest of the constants is straightforward. The constants with subscripts denote the leaf nodes of Clements' tree and the atomic feature matrices.

Pierrehumbert & Beckman's (1988:119) partial prosodic structure can be expressed as follows:

$$(2-39) \quad \left[\text{SYL} \left\langle \left[\begin{array}{cc} \text{ONSET} & \langle s \rangle \\ \text{MORA} & \langle e \ e \rangle \end{array} \right] \left[\begin{array}{cc} \text{ONSET} & \langle t \rangle \\ \text{MORA} & \langle a \ a \rangle \end{array} \right] \left[\begin{array}{cc} \text{ONSET} & \langle w \rangle \\ \text{MORA} & \langle a \rangle \end{array} \right] \right\rangle \right]$$

This matrix introduces a new piece of notation, the angle bracket. Angle brackets are used in the representation of sequences. The word *seetaawa* consists of a sequence of three syllables. Each syllable in turn contains an ONSET sequence and a MORA sequence. As there are no complex onsets the onset sequences have only one element. The fact that onsets occur before moras in a syllable may be expressed using the following constraint:

(2-40) **Onsets Precede Moras:**

$$\forall xyz \text{ onset}(x,y) \wedge \text{mora}(x,z) \rightarrow y \prec z$$

The formula describing the above feature matrix (and also describing Pierrehumbert & Beckman's original tree) appears in (2-41).

$$\begin{aligned} (2-41) \quad & \text{word}(a) \wedge \\ & \text{syl}(a,b) \wedge \\ & \quad \text{onset}(b,b_1) \wedge s(b_1) \wedge \\ & \quad \text{mora}(b,b_2) \wedge e(b_2) \wedge \\ & \quad \text{mora}(b,b_3) \wedge e(b_3) \wedge b_2 \prec b_3 \wedge \\ & \text{syl}(a,c) \wedge \\ & \quad \text{onset}(c,c_1) \wedge t(c_1) \wedge \\ & \quad \text{mora}(c,c_2) \wedge a(c_2) \wedge \\ & \quad \text{mora}(c,c_3) \wedge a(c_3) \wedge c_2 \prec c_3 \wedge \\ & \text{syl}(a,d) \wedge \\ & \quad \text{onset}(d,d_1) \wedge w(d_1) \wedge \\ & \quad \text{mora}(d,d_2) \wedge a(d_2) \wedge \\ & b \prec c \prec d \end{aligned}$$

As a final example, consider a structure involving re-entrancy, such as that proposed by McCarthy & Prince (1989), reproduced above, for the word *kattab*. It may be represented in a feature matrix as follows.

$$(2-42) \quad \left[\text{SYL} \left\langle \left[\begin{array}{cc} \text{ONSET} & \langle k \rangle \\ \text{MORA} & \langle a \ t^{\boxed{1}} \rangle \end{array} \right] \left[\begin{array}{cc} \text{ONSET} & \langle t^{\boxed{1}} \rangle \\ \text{MORA} & \langle a \ b \rangle \end{array} \right] \right\rangle \right]$$

The boxed index $\boxed{1}$ indicates that the two t 's are in fact just one token. A special notation is required here because of the background assumption that multiple mention of a symbol in a diagram is taken to imply the existence of equally many corresponding objects in the domain. It is therefore only necessary to mention the t once, as in (2-43), and it does not matter which place it is mentioned in.

$$(2-43) \quad \left[\text{SYL} \left\langle \left[\begin{array}{cc} \text{ONSET} & \langle k \rangle \\ \text{MORA} & \langle a \ t^{\boxed{1}} \rangle \end{array} \right] \left[\begin{array}{cc} \text{ONSET} & \langle \boxed{1} \rangle \\ \text{MORA} & \langle a \ b \rangle \end{array} \right] \right\rangle \right]$$

Now that we have sequences, it is necessary to be able to depict constraints on their alignment. The standard notation for autosegmental association will be carried over directly into the feature matrix notation. The overlap statements in (2-44a) will be depicted using lines as in (2-44b).

(2-44) a. $a \circ d \wedge b \circ d \wedge c \circ e \wedge c \circ f$

b.
$$\left[\begin{array}{c} F \\ G \end{array} \begin{array}{c} \langle a' b' c' \rangle \\ \vee \wedge \\ \langle d' e' f' \rangle \end{array} \right]$$

2.6 Some Abbreviatory Conventions

The formulae provided above tend to be somewhat long-winded. From a programming languages perspective, it is natural to view these formulae as comprising a low-level ‘machine language’. We can employ a variety of ‘higher level’ constructs, just so long as they can be compiled away in a reasonable number of steps. This allows descriptions to be more succinct but no less precise. For the sake of clarity these are defined by example. Their presentation is in no particular order.

The first abbreviatory device is for stating appropriateness constraints. The unabbreviated formula in (2-45a) is written as (2-45b).

(2-45) **Abbreviatory Device 1:**

a. $\forall xy \text{ place}(x) \wedge x \delta y \rightarrow \text{coronal}(y) \vee \text{anterior}(y) \vee \text{distrib}(y)$

b. $\text{place} \Rightarrow [\text{coronal}, \text{anterior}, \text{distrib}]$

If it is necessary to constrain the amount of branching which is permitted from nodes of a particular sort to nodes of some other sort, the following notation will be used, where n is the minimum and m is the maximum (cf. 2-18,19). The example given in (2-46) states that syllables consist of one or two moras.

(2-46) **Abbreviatory Device 1':**

a. $\forall xy (\text{syl}(x) \wedge x \delta y \rightarrow \text{mora}(y)) \wedge$
 $\forall xy_1 y_2 y_3 (\text{syl}(x) \wedge x \delta y_1 \wedge x \delta y_2 \wedge x \delta y_3 \rightarrow y_1 = y_2 \vee y_1 = y_3 \vee y_2 = y_3) \wedge$
 $\forall x (\text{syl}(x) \rightarrow \exists y x \delta y)$

b. $\text{syl} \Rightarrow [\text{mora}(1,2)]$

A related device has already been proposed in section 5 for relating the sort predicates with binary relations which were partitions of δ . It is reformulated as (2-47):

(2-47) **Abbreviatory Device 2:**

a. $x \delta y \wedge s(y)$

b. $s(x,y)$

Another kind of appropriateness constraint is expressed in (2–48a). Its abbreviated form is in (2–48b).

(2–48) **Abbreviatory Device 3:**

- a. $\forall x \text{ nasal}(x) \equiv +\text{nasal}(x) \vee -\text{nasal}(x)$
 $\forall x +\text{nasal}(x) \rightarrow \neg -\text{nasal}(x)$
- b. $\text{nasal} = \{ +\text{nasal}, -\text{nasal} \}$

The fourth device is for composing relations.

(2–49) **Abbreviatory Device 4:**

- a. $f(a,b) \wedge g(b,c)$
- b. $f|g(a, c)$

In general, $f|g(x, z)$ is true just in case there is some y such that $f(x,y)$ and $g(y,z)$ ¹⁰.

The next device is for sequences. A similar device will be assumed for sets, using ‘{’ and ‘}’ instead of ‘⟨’ and ‘⟩’.

(2–50) **Abbreviatory Device 5:**

- a. $f(a,b) \wedge f(a,c) \wedge f(a,d) \wedge b \prec c \prec d$
- b. $f(a, \langle b, c, d \rangle)$

Sometimes we require that a certain sequence description be exhaustive. Again, a similar device will be assumed for sets.

(2–51) **Abbreviatory Device 6:**

- a. $f(a,b) \wedge f(a,c) \wedge f(a,d) \wedge b \prec c \prec d \wedge$
 $\forall x (f(a,x) \rightarrow x = b \vee x = c \vee x = d)$
- b. $f(a, \langle\langle b, c, d \rangle\rangle)$

Note that the above conventions primarily concern the δ and \prec relations. Of course, the \circ relation is closely related to the δ relation, insofar as dominance implies overlap. However, there will be no abbreviatory conventions adopted for overlap statements above and beyond those which come from dominance. This is because its role in linguistic descriptions appears to be less central than that played by δ and \prec .

So far in this chapter we have seen a number of notational conventions from the phonological literature and have explored their formal interpretation. Along the way, a description language \mathcal{L} is being built up. This language will need to encompass

¹⁰ These composite relations are instances of binary relations which are not refinements of the immediate dominance relation, in contrast with the binary relations resulting from the second abbreviatory device.

not only representations but also phonological rules and principles. Once these have been discussed (in the next section), a formal definition of \mathcal{L} can be given (the following section).

2.7 Rules and Principles

Although rules are considered to play a central part in phonological description, the general conception of a phonological rule is even less clear than the general conception of a phonological representation. In autosegmental phonology some have simply incorporated the autosegmental representation notation into the SPE rule notation (e.g. Hall 1989). Goldsmith (1990) catalogues a bewildering variety of attested operations (see Bird and Ladd *forthcoming* for a review). Others have complained that "standard autosegmental approaches to rules fail because they allow too many unattested rule types" (Archangeli & Pulleyblank 1987:32).

The constraint on rules which arises in a monostratal framework is that the arrow (i.e. implication) relates classes of objects which are at the same level (i.e. stated in the same language) rather than on different levels. For example, consider the simple rule in (2-52a), which is translated into the present formalism as (2-52b). A paraphrase of the rule appears in (2-52c)¹¹. This is just the standard semantics for implication.

- (2-52) a. $[-\text{voice}] \rightarrow [-\text{nasal}]$
 b. $\forall x \text{ } -\text{voice}(x) \rightarrow -\text{nasal}(x)$
 c. If a segment is voiceless then it is also not nasal

This kind of interpretation of a phonological rule is equivalent to viewing the rule as a partial description. For the above example, the description is as follows: "all segments are either voiced or non-nasal". This description is perhaps slightly unusual for phonology in that it employs negation and disjunction¹². However, it is natural to admit such possibilities in contexts where the description/object distinction is made.

¹¹ This interpretation of the arrow was made explicit in some early generative writings. For example, Schane (1973:36) states that the arrow "is to be read as 'is also' or 'implies'."

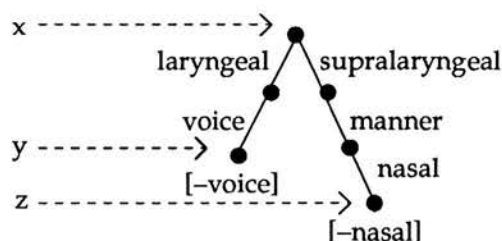
¹² Note that Hudson's (1980) treatment of automatic alternations effectively employs disjunctive phonological descriptions.



Given this conflation of rules and representations, the term we adopt which covers both is ‘constraint’.

The constraint in (2–52b) does not involve predicates other than the sort predicates. Since our descriptions are hierarchically structured, it is necessary to show how the δ relation is incorporated into constraints. In a hierarchical structure, the two features mentioned in the above rule may be linked to different parts of a tree structure. For example, in Clements’ hierarchically structured tree, the features voiced and nasal appear under different class nodes. As our descriptions are generally only partial (and so may not be fully specified) the reference to a node which is low in the ‘tree’ structure necessarily assumes the existence of the intervening nodes¹³. However, recall that our composite relation notation is ultimately an abbreviation for an expression which has existential quantifiers. Therefore it is only necessary to explicitly state the existence of the leaf node (here, z).

- (2–53) a. $\forall xy \text{ laryngeal}|\text{voice}(x,y) \wedge \neg\text{voice}(y) \rightarrow \exists z \text{ supra.l}|\text{manner}|\text{nasal}(x,z) \wedge \neg\text{nasal}(z)$
 b.



For some time it has been argued that default rules should play a part in phonological descriptions (e.g. Stanley 1967, Chomsky & Halle 1968:382ff). The generally accepted interpretation of a default rule such as that in (2–54a) is given in (2–54b).

- (2–54) a. $[+low] \rightarrow [+back]$
 b. If a segment is specified as +low and it is not specified as –back then assume that it is +back

Generalizing from this, the interpretation of a default rule $a \rightarrow b$ is as follows. If

¹³ Archangeli & Pulleyblank (pers. comm.) have proposed a ‘Node Generation’ convention, which is stated as follows: “a rule or convention assigning some feature or node x to some node b creates a path from x to b ”. A similar proposal appears in (Avery & Rice 1989:183), called the ‘Node Activation Condition’.

a segment s meets the description a and it is consistent to assume that b is true of s , then take b to be true.

The difference in interpretation between standard rules and default rules is not signalled directly in the notation. An attractive way to signal this difference formally is to employ Reiter's (1980) notation. Reiter considers that "inferences sanctioned by default are best viewed as beliefs which may well be modified or rejected by subsequent observations". Although his approach concerns the reasoning about beliefs, there is no reason to suppose that his approach cannot be equally well applied to other areas of default behaviour. Reiter's approach involves enriching a first order theory with a collection of metarules which employ a special operator M . $M(\phi)$ is read as "it is consistent to assume ϕ ". The default rule (2-55a) is expressed as the formula in (2-55b).

- (2-55) a. $[+low] \rightarrow [+back]$
 b. $\forall x +low(x) \wedge M +back(x) \rightarrow +back(x)$

Properties which are *required* to be present for a rule to apply are specified in the structural description (i.e. rule left hand side) as normal. Properties which only need to be *compatible* with the context are prefixed by the M operator. To guarantee a default theory has an extension, Reiter advocates the adoption of rules in the following format, where w is any formula, possibly involving quantifiers, whose free variables are x (a vector). The expression in (2-56a) is Reiter's format, and the one in (2-56b) is our abbreviated version.

- (2-56) a. $\alpha(x) \wedge M w(x) \rightarrow w(x)$
 b. $\alpha(x) \xrightarrow{d} w(x)$

Rather conveniently, Reiter's restricted format includes the expressions required for default rules in phonology. This notation can be extended to encompass hierarchical structure, as there is no constraint on (2-56) involving unary predicates only.

This provides the expressive power to be able to distinguish the default value of a feature from the default existence of a node. It also permits the statement of defaults such as the following:

- (2-57) a. $\forall x \text{ syl}(x) \xrightarrow{d} \exists yz (y \neq z \wedge \text{mora}(x, \langle y z \rangle))$
 b. By default, syllables consist of two moras

There are other ways to approach the matter of defaults, of which Reiter's is just one. Another is the DATR system of Evans and Gazdar (1990) which has been successfully applied to a range of natural language phenomena. Another is to employ the compilation strategy discussed in chapter 1 section 2. The discussion this matter deserves goes beyond the scope of the present study.

2.8 The Language \mathcal{L}

In this section a classical first-order theory will be defined. It is classical because the double negation of a formula is always equivalent to the formula. It is first-order because quantification is over individual variables and not predicates. The theory is also function-free, in that the only arguments a predicate can take are variables.

We begin by defining the primitive symbols of \mathcal{L} :

1. A set of individual variables \mathcal{V}
2. A finite set of sort symbols \mathcal{S}
3. The symbols $\delta, \delta^*, <, \preceq, \circ, =, \neg, \vee, \wedge, \forall, \exists, \rightarrow, \equiv, (,), \top, \perp$

The pair $(\mathcal{V}, \mathcal{S})$ form the signature of \mathcal{L} . Modifying \mathcal{V} and \mathcal{S} is how we 'customize' \mathcal{L} for a particular phonological theory. However, we shall normally only be explicit about \mathcal{S} in this connection.

The TERMS, or denoting expressions, are just the individual variables. The FORMULAE (wffs) of \mathcal{L} are constructed out of the primitive symbols as follows.

1. An expression consisting of a sort symbol followed by an individual variable is a wff
2. $x \delta y, x \delta^* y, x < y$ and $x \circ y$ are wffs for all individual variables x, y
3. If ϕ is a wff then so is $\neg\phi$
4. If ϕ and ψ are wffs, so are $(\phi \vee \psi), (\phi \wedge \psi), (\phi \rightarrow \psi), (\phi \equiv \psi)$

5. If ϕ is a wff, then for each individual variable x , $(\forall x)\phi$ and $(\exists x)\phi$ are wffs

The above will be viewed as a formal definition of the syntax of a phonological description language \mathcal{L} . A formal semantics for this language is inherited from the standard model theoretic semantics of classical first-order predicate calculus. Briefly, we assume the existence of a domain of individuals \mathcal{D} and a valuation function assigning a member of \mathcal{D} to each member of \mathcal{V} . In the present context, \mathcal{D} will usually be a set of nodes. However, we shall ultimately want to think of \mathcal{D} as being a set of utterances. The sort predicates range over properties of objects and those of degree two range over binary relations.

The remainder of this section is devoted to the statement of various constraints on formulae.

The first group concern the relationship between the various standard connectives. Further axioms for equality substitution are omitted here.

- (2-58) 1. $\phi \wedge \psi$ iff $\neg(\neg\phi \vee \neg\psi)$
 2. $\phi \rightarrow \psi$ iff $(\neg\phi) \vee \psi$
 3. $\phi \equiv \psi$ iff $(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$
 4. $\exists x\phi$ iff $\neg(\forall x)\neg\phi$

The two precedence relations interact according to (2-59a). The two dominance relations interact according to (2-59b).

- (2-59) a. $\forall xy \ x \preceq y \equiv x \prec y \vee x = y$
 b. $\forall xy \ x \delta^* y \equiv x \delta y \vee \exists z (x \delta z \wedge z \delta^* y)$

The \circ , \prec and δ^* relations must satisfy the following axioms.

Temporal Axioms:

- (2–60) a. **Overlap is reflexive:**
 $\forall x \ x \circ x$
- b. **Overlap is symmetric:**
 $\forall xy \ x \circ y \rightarrow y \circ x$
- c. **Precedence is asymmetric:**
 $\forall xy \ x \prec y \rightarrow \neg y \prec x$
- d. **Precedence is disjoint from overlap:**
 $\forall xy \ x \prec y \rightarrow \neg x \circ y$
- e. **Precedence is transitive (through overlap):**
 $\forall wxyz \ w \prec x \wedge x \circ y \wedge y \prec z \rightarrow w \prec z$
- f. **Time is linear:**
 $\forall xy \ x \prec y \vee x \circ y \vee x \succ y$
- g. **Temporal locality in dominance:**
 $\forall xy \ x \delta^* y \rightarrow x \circ y$

Constraints on constituents:

- (2–61) a. **‘Set values’ have a linear ordering in $\prec, =, \succ$ (stated for all s):**
 $\forall xyz \ x \delta y \wedge x \delta z \wedge s(y) \wedge s(z) \rightarrow y \prec z \vee y = z \vee y \succ z$
- b. **Unary values (stated for some s):**
 $\forall xyz \ x \delta y \wedge x \delta z \wedge s(y) \wedge s(z) \rightarrow y = z$

At the start of this section we saw two sets \mathcal{V} and \mathcal{S} , which are the variables and the sort symbols of \mathcal{L} . These are parameters, and changing them results in a different language. The collection of such parameters (here $\langle \mathcal{V}, \mathcal{S} \rangle$) is known as a *signature*. As it stands, \mathcal{L} has a large collection of axioms which are essentially arbitrary. As these are refined and it becomes clearer what the structure of the intended models is, a more revealing approach will be possible. Rather than parameterizing \mathcal{L} with two sets, we can parameterize it with algebraic structures. This suggestion is taken up in (Bird 1991b)¹⁴.

2.9 Combining Descriptions

As already mentioned, the lexicon is the locus of morphologically simplex signs. The derivation of a morphologically complex expression involves the combination of sim-

¹⁴ The default operator M and the default arrow \xrightarrow{d} are not a part of \mathcal{L} but belong to a metalanguage, \mathcal{L}' . This metalanguage—and not \mathcal{L} —is the language in which linguistic descriptions will be couched. However, I shall ignore the distinction between \mathcal{L} and \mathcal{L}' because the precise formulation of defaults is not central to this study.

pler signs. In particular, the combination respects the internal structure of signs along the phonology, syntax and semantics dimensions (Oehrle 1988). The sole combination operation for phonological descriptions involves a mixture of conjunction and equality. Although this view of phonological combination may seem impoverished its expressive power is considerable, as this section endeavours to show.

A collection of phonological constraints ϕ is, in some sense, ultimately about the 'root' node x of a phonological structure (whether or not this exists lexically). This so-called root node provides an important handle on the structure, as will be seen below. I shall assume that the phonological descriptions in different lexical entries do not share variables. When morphemes are brought together, there must be some way to equate at least one pair of variables across the two separate structures, in order to express the fact that they are (now) both part of a single description. For this reason, a collection of constraints ϕ with root node x will sometimes be written $\phi(x)$, or equivalently $\langle x, \phi \rangle$. The phonology attribute of a sign is such an ordered pair. The simplest way to combine two such attributes is as in (2-62); their roots are equated and their formulae are conjoined.

$$(2-62) \quad \langle x, \phi \rangle . \langle y, \psi \rangle = \langle x, x = y \wedge \phi \wedge \psi \rangle$$

However, it will often be the case that the phonological dimension of one morpheme μ_1 will be embedded in that of another morpheme μ_2 . In this case, the root node of μ_1 will not be equated with that of μ_2 but rather with some constituent of the root node of μ_2 . Here the simpler type of combination will be employed to illustrate the expressive capabilities of this general approach to phonological combination.

First, consider the two set descriptions in (2-63a). (If these were part of a lexical entry then the variables would all be existentially quantified.)

$$(2-63) \quad \begin{array}{ll} \text{a.} & \langle x, f(x, \{y_1 \cdots y_n\}) \wedge \forall_i \neq_j y_i \neq y_j \rangle \\ \text{b.} & \langle x', f(x', \{y'_1 \cdots y'_{n'}\}) \wedge \forall_i \neq_j y'_i \neq y'_j \rangle \end{array}$$

Combining these according to (2-62) produces (2-64).

$$(2-64) \quad \langle x, x = x' \wedge f(x, \{y_1 \cdots y_n y'_1 \cdots y'_{n'}\}) \wedge \forall_i \neq_j y_i \neq y_j \wedge y'_i \neq y'_j \rangle$$

Although the y_i are distinct and y_i' are distinct, there is nothing preventing $y_i = y_j'$ for some i, j . (Even more options are possible if the y_i and the y_i' are not required to be distinct, cf. Pollard & Sag 1987:47.) Only three options are of direct relevance here, and all three assume that the individual conjuncts specify sequences (and not sets¹⁵).

Suppose that the individual sequences are disjoint. This is reasonable given the assumption that the variables in each lexical entry are disjoint from those in all other entries. Now the conjunction of sequences corresponds to SEQUENCE UNION, exemplified in (2-65).

$$(2-65) \quad f(v, \langle w x \rangle) \wedge f(v, \langle y z \rangle) \equiv f(v, \langle w x y z \rangle) \vee f(v, \langle w y x z \rangle) \vee f(v, \langle y w x z \rangle) \vee f(v, \langle w y z x \rangle) \vee f(v, \langle y w z x \rangle) \vee f(v, \langle y z w x \rangle)$$

If we add the stipulation that $x \prec y$, then the result corresponds to CONCATENATION:

$$(2-66) \quad f(v, \langle w x \rangle) \wedge f(v, \langle y z \rangle) \wedge x \prec y \equiv f(v, \langle w x y z \rangle)$$

If the disjointness condition is dropped and exhaustiveness conditions are added to both conjuncts, then their conjunction corresponds to SEQUENCE UNIFICATION.

$$(2-67) \quad f(v, \langle\langle w x \rangle\rangle) \wedge f(v, \langle\langle y z \rangle\rangle) \rightarrow w = y \wedge x = z$$

2.10 Conclusion

In this chapter it has been shown how the diagrammatic conventions adopted in phonology can be studied by elucidating the properties of the objects and relations which are encoded in these diagrams. The properties of objects were expressed in terms of sorts, and sorts were shown to be organized into a lattice. A number of hierarchical and temporal relations were defined and many aspects of their interaction were studied. An important result is the elucidation of the relationship between dominance and overlap¹⁶. Next it was shown how the resulting system bears close

¹⁵ Although I do not employ set descriptions, they are presented here because of their heavy use in other areas of grammar (e.g. Pollard & Sag 1987).

¹⁶ These are normally conflated in phonological theorizing, but have been distinguished by Sagey (1986:20ff) and Hayes (1990). However, Sagey's suggestion is neither explicit nor exemplified, and Hayes' suggestion—which does not have these shortcomings—nevertheless suffers from a number of flaws as discussed in Appendix B.

resemblances with the feature matrix notation. Finally, the language \mathcal{L} was defined and it was shown how phonological descriptions could be combined.

Throughout this chapter I have tried to maintain a clear distinction between descriptions and objects. The objects are the graph and matrix diagrams, whereas the descriptions are the members of \mathcal{L} . However, the members of \mathcal{L} ultimately describe real-world utterances. On this view, graphical notations stand for classes of models, or for minimal models. Feature matrices are not partial descriptions but complete objects. Johnson (1990:178) has pointed out that the first-order approach to featural organization has significant advantages over the so-called 'designer logic' approach, in that important properties such as soundness, completeness, decidability and compactness do not need to be proven from scratch (and reproven each time a modification is made), and there exists a significant body of work on satisfiability algorithms for first-order formulae.

In proposing \mathcal{L} , what might be called *temporal feature logic*, I have tried to draw simultaneously on work in temporal logic and work in feature logic. The result is not particularly startling or rich in its present form (at least from the point of view of theoretical logic), but it does point towards a potentially fruitful area of interaction which is yet to be addressed from a theoretical standpoint. However, the present concerns are purely practical, and we now have more than enough machinery to enable the further elucidation of extant phonological proposals.

In the next chapter a detailed application of this logical framework is described.

Chapter 3

A Theory of Phonological Structure

For a long time it was assumed that the smallest unit which needed to be recognized in phonology was the phoneme. Swadesh, drawing on work by Bloomfield, Jones, Sapir and others, provided the following compact statement about the phoneme:

The phonemic principle is that there are in each language a limited number of *elemental* types of speech sounds, called phonemes, peculiar to that language; that all sounds produced in the employment of the given language are referable to its set of phonemes; that only its own phonemes are at all significant in the given language. (Swadesh 1934:32, emphasis added)

Once phonemes were characterized according to the oppositions they entered into, it became clear that these oppositions were not arbitrary but revealed the organization of phonemes into natural classes. These classes had defining properties, such as voicing or lip rounding, which could be cast in terms of distinctive features (Trubetzkoy, Jakobson and others). A phoneme (or a segment) was viewed as a set of features, and an utterance corresponded to a sequence of these sets, or a two-dimensional array of features. Although features are properties of phonemes, they could also be regarded as entities in their own right which come together to form phonemes, a duality which still exists today. Chomsky & Halle (1968:299ff) showed how these features do not all have the same status but can be grouped; there are major class, cavity, manner, source and prosodic features. Some of these groups have subgroups in turn. However the purpose of this grouping was only for exposition. Clements

(1985) and others have argued in favour of hierarchical representations of segments which reflect this grouping (examples were given in chapter two). Phonological observations could then be stated more succinctly and we had the beginnings of an explanation of why certain groups of features pattern together for the purposes of assimilation and other phenomena (see McCarthy 1988 for a detailed overview). Sagey (1986) provided a detailed study in support of the claim that the feature hierarchy is based on phonetics. "Features are grouped according to the articulator in the vocal tract that they are executed by. Articulators are grouped according to their acoustic effects on the formant structure" (Sagey 1986:2). Browman & Goldstein (1989) further explored this phonetic basis of feature hierarchy, making it more explicitly phonetic and providing a more plausible model of the causal relation between articulatory and acoustic properties.

If we survey this particular strand of development in detail it becomes clear that the lasting and non-controversial categories (or groupings) employed by phonology are those which also have a strong phonetic basis. Those without this basis (such as Clements' 'manner' and 'supralaryngeal' nodes) have provoked a deal of controversy, about which no consensus has yet been reached. Indeed, the empirical arguments concerning such phonetically unmotivated categories often conflict, as will be discussed in section 1. Therefore, I believe we can observe a persistent historical trend towards the view that the basic units recognized by phonology correspond directly to the articulatory parameters which speakers have direct control over. This should not be surprising. On the contrary, paraphrasing Fowler's (1980:118–9) arguments, it would be surprising if a category arose in evolution or ontogeny which did not bear a close resemblance to the actualization with which it co-evolved or co-developed; furthermore, such a category (if it existed) would vastly complicate the communication process.

In this chapter I shall attempt to take this trend further towards its logical conclusion. The first section takes a look at the phonological evidence for the hierarchical organization of features, and the second section proposes an articulatory model and shows how acoustic properties can be viewed as emergent. The next section shows

how this model can be linked to a model of prosodic structure. The fourth section formalizes the proposals, and the fifth section discusses the emerging view of phonological organization.

3.1 The Evidence for Hierarchical Organization

There have been numerous proposals for the hierarchical organization of features in the literature (see McCarthy 1988 and Broe *to appear* for surveys). Perhaps the earliest clear statement was made by Lass:

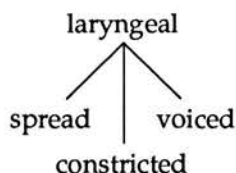
- (i) Any phonological/phonetic segment is represented as a two-part matrix, consisting of submatrices labelled [oral] and [laryngeal].
 - (ii) The notational independence of the two parameters implies that each is a possible proper domain for a phonological rule: (*sic*) in addition to the whole segment being such a domain.
- ... submatrices as wholes can function in rules; not only in deletions, but also in rules appealing to the notions 'homorganic' and 'identical'. (Lass 1976:154–5)

Lass' identification of the *functions* of his submatrices is still widely assumed today. For example, Yip states that "three kinds of evidence can be used to argue for a particular constituent structure for distinctive features: (i) constituents spread as units; (ii) constituents delete/detach as units; (iii) constituents are identified as units by rules which compute identity, such as the obligatory contour principle" (1989b:349).

For the remainder of this section a number of feature groupings which have been proposed in the literature are presented and evaluated. When discussing a particular feature grouping (such as 'supralaryngeal') the potential sub-groupings (such as 'place' and 'manner') are not relevant; it is sufficient to think of each grouping as a set in isolation from any further structure on that set.

The Laryngeal Node. Clements (1985) groups the features [spread], [constricted] and [voiced] under the laryngeal node, which we may depict schematically as in (3–1).

(3-1) Clements' Laryngeal Node:

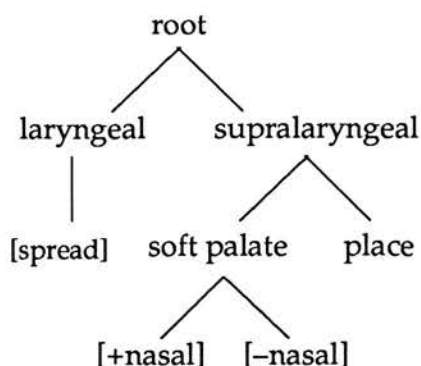


In support of this he cites phenomena from Thai, Klamath (Penutian, North America) and Proto-Indo-Iranian, which indicate that the features thus grouped behave as a unit in phonological processes. Syllable initial stop consonants in Thai can be voiced, voiceless aspirated and voiceless unaspirated. Syllable final stops can only be voiceless unreleased. Although Clements is not explicit about the interpretation of this data, McCarthy (1988:90) explains that “delinking of the laryngeal node in syllable-final position reduces all three categories to just the unmarked one”. Presumably a similar account is given for Klamath, where the distinction between voiced, voiceless and glottalised obstruents is neutralised when the obstruents immediately precede a stop. For both Thai and Klamath, Clements assumes that the laryngeal sub-trees removed from these segments are ultimately replaced by a structure containing the default values of these features. In a similar vein, Sagey (1986:35) cites the case of aspirated voiceless stops in Kinyarwanda (Bantu, Rwanda & Zaïre) which alternate with voiceless or aspirated nasals, as the following examples illustrate:

	/in-papuro/	[im _h hapuro]	“paper”
(3-2)	/n-toora/	[n _h hoora]	“vote for me”, “I vote”
	/in-ka/	[i _h ha]	“cow”

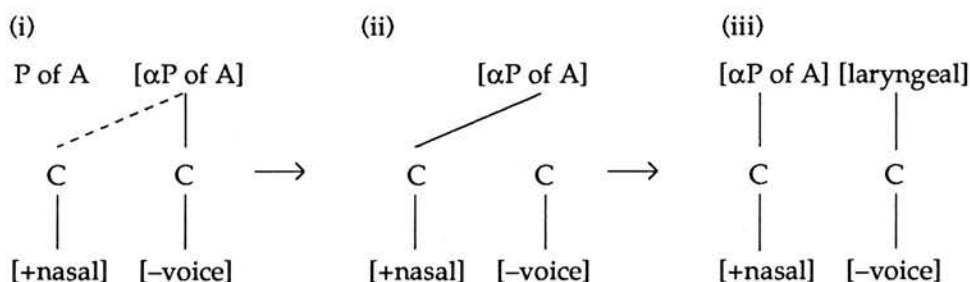
Here the circle indicates voicelessness and the h indicates aspiration. Sagey claims (presumably on the basis of uncited evidence) that these Nh ‘clusters’ are actually prenasalized stops at some level. She states: “since we represent the aspiration of the oral stops on an independent, unordered tier (the laryngeal tier), this realization of the aspiration of the oral stop on the added nasal portion is predicted, given that prenasalization is a merging of [+nasal] into the root node of the stop, resulting in the structure in [(3-3)].”¹

¹ Sagey’s analysis is only half of the story. We need to transfer the place specification from the (underlying) stop to the nasal, leaving the stop without an oral closure. However, this alignment of velic and oral gestures is clearly not expressible in the above tree notation. The only way it can be



Sagey's use of this data as an argument for the laryngeal class node is weakened by the existence of an analysis of the data² which does not require this node. A similar example comes from KiRundi (actually a dialect of Kinyarwanda from Burundi) where "a nasal is always homorganic to a following consonant, and a voiceless consonant following a nasal loses its oral point of articulation, leaving only its aspirate character behind, what we transcribe as an *h* in syllable onset position" (Goldsmith 1990:283). Goldsmith does not employ a laryngeal class node but instead views laryngeal as a place of articulation, or 'P of A' (cf. Sagey, for whom places of articulation can only be oral).

(3-4) Goldsmith's (1990:284) Analysis for KiRundi:



Therefore Sagey's conclusion from the Kinyarwanda data about the necessity of a laryngeal class node rests on the implicit assumption that the laryngeal place of articulation does not form a part of the place node. In fact, Sagey's place node only concerns oral places of articulation, and it might be more appropriately labelled

expressed is to use a structure with two root nodes, where the first is the nasal (with laryngeal and place features from the stop) and the second is the 'stop' itself (minus the place features). However, Sagey has opted for a structure with only one root node.

² Goldsmith's analysis (below) is for the language KiRundi (Burundi). Note that Kinyarwanda and KiRundi are dialects of the same language.

'oral cavity' rather than 'place'. The argument from Kinyarwanda/KiRundi for the laryngeal node is not sufficiently independent to be particularly compelling.

Another example is from Classical Greek (McCarthy 1988:90), where stop clusters regressively assimilate in both voicing and aspiration. McCarthy's conclusion is that the laryngeal node spreads. However, as no data is given it is difficult to evaluate this claim.

Even if one took this collection of evidence for the laryngeal node as compelling, it only motivates a hierarchical structure in conjunction with similar evidence for the independent behaviour of the laryngeal features themselves (such as voiced, spread and constricted). Given the absence of such arguments we might as well have just one polyvalent laryngeal feature (cf. McCarthy 1988:94ff, Goldsmith 1990:293) or a bundle of binary features which may only be treated as a group. (If one rejected the above evidence for the laryngeal node, such a polyvalent feature or feature bundle could be sited inside the root node, to indicate that it only spreads in conjunction with the root node itself, as McCarthy has proposed for certain manner features.) In none of these cases is there a laryngeal node which further dominates other nodes.

The only case for the independent functioning of subsets of the laryngeal set concerns tone features. Although the autosegmental approach has its origins in analyses of tone, there has been relatively little discussion of what the relevant distinctive features are (see Wang 1967, Pulleyblank 1986:125, Yip 1989a for suggestions). However, it is clear from the vast array of analyses involving the spreading, delinking and obligatory contour principle (OCP) effects of tone that the tone features may operate independently of the other laryngeal features. For example, it is uncommon for the spreading of tones to non-adjacent vowels to be hindered by the presence of particular intervening consonants (see Kisseberth 1984:137,152 for a rare exception from Digo). Whatever these tone features are, there have been no attempts to locate them in the feature hierarchy (Goldsmith 1990:293).

An explanation for the apparent difficulty in treating the laryngeal features comes again from phonetics. The larynx is an unusual member of the vocal tract, being

involved in initiation, phonation and articulation. The glottis provides initiation (i.e. air pressure) for implosives and ejectives. For example, an ejective is produced with glottal and oral closure; when the larynx (which houses the glottis) is raised, the air trapped between the two closures is compressed and when the oral closure is released this air escapes, with an audible result. The glottis is also involved in phonation, being responsible for voicelessness, whispery voice or murmur, voicing and creaky voice (see Ladefoged 1971:17 for an identification of nine points on this continuum). In any of the voicing states, the observed pitch is a function of vocal fold tension as well as subglottal air pressure (which is pulmonic). Finally, the larynx has an articulatory function (e.g. the glottal stop and perhaps the epiglottal stop). These three functions³ generally do not co-occur, although there are some exceptions. For example, for voiced implosives the downward movement of the larynx is the initiator and the glottis provides phonation. It is therefore possible for there to be a voicing distinction for implosives (as Ladefoged et al. 1976 have observed in Owerri Igbo) but not for ejectives (which are voiceless). Existing phonological accounts of laryngeal features are unable to express naturally (if at all) the cooccurrence patterns of initiatory, phonatory and articulatory functions of the larynx. In the light of this extraordinary complexity of the larynx it is not surprising that the above discussion of the laryngeal node was inconclusive.

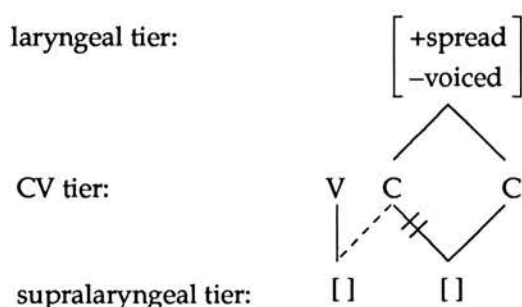
The Supralaryngeal Node. Lass (1976) discusses English consonant reduction to the glottal stop, providing evidence from New York and Scots dialects (pp. 149-50). He analyses these by characterizing “every segment ... as (at least) ‘bi-gestural’: there are two relevant articulatory configurations, one laryngeal and the other supralaryngeal. ... Thus [ʔ] and [h] are defective; their matrices lack defining specifications for features that are purely intra-oral, like ‘coronal’, ‘back’ and so forth. They are missing an entire component or parameter that is present in ‘normal’ segments” (p. 153). Consonant reduction is then just the deletion of the supralaryngeal gesture from a consonant, concurrently replacing the supralaryngeal closure with a laryngeal one (p. 155). (Note that Lass’ discussion does not settle the question of the representation

³ See Catford 1988:23ff, 51ff, 100ff for detailed discussion of initiatory, phonatory and articulatory uses of the glottis/larynx.

of [h]; see Iverson 1989 for further discussion.)

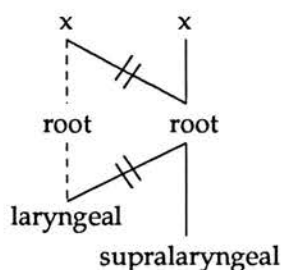
Evidence from Icelandic aspiration has received much attention in the literature on feature hierarchy (Thráinsson 1978, Clements & Keyser 1983:79, Clements 1985:233-4, Árnason 1986, Sagey 1986:32ff, Iverson 1989, Hayes 1990). Thráinsson's basic observation is that underlying geminate voiceless stops appear as preaspirated: /pp, tt, kk/ → [hp, ht, hk]. Thráinsson analyses this as involving the spreading of a supralaryngeal node of a preceding vowel onto the first position occupied by the geminate consonant, along with the delinking of the supralaryngeal specification for that consonant.

(3-5) Thráinsson's Analysis:



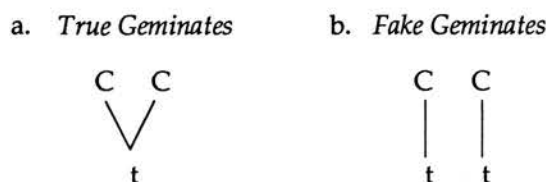
It is as though the supralaryngeal specification of the vowel has extended rightwards, pushing the supralaryngeal specification for the consonant further right. However, Árnason (1986:19) takes the opposing view that “we are dealing with a movement of the openness of the glottis connected with the stop towards the nucleus. It is thus an anticipatory opening of the glottis ... [which occurs] when a vowel preceded that did not become long in the [historical] quantity shift.” Similarly Sagey’s account involves the leftwards realignment of the laryngeal node that was initially part of the geminate. In (3-6), a new root node needs to be created to take the newly created association lines.

(3-6) Sagey's Analysis:



Now we have seen two analyses. In the first, a *supralaryngeal* node 'moves rightwards', and in the second, a *laryngeal* node 'moves leftwards'. Not only do the analyses conflict, the generalization that preaspiration affects only geminates (assumed by Clements, Sagey and Hayes following Thráinsson) is controversial. Árnason (1986:13) cites examples such as /opna/ where a lengthening rule is required to produce /oppna/ before the preaspiration rule can apply. However, Árnason states that the evidence for such a rule is not clear. Furthermore, the very motivation for the existence of preaspiration in tautomorphemic cases comes from the heteromorphemic ones. However, the latter do not necessarily motivate the former, for there are many instances of phonological process which affect heteromorphemic geminates⁴ without affecting tautomorphemic geminates⁵ (Hayes 1986a:490, Hayes 1986b:341, McCarthy 1986:218). The structural distinction between the two kinds is a result of the OCP, which only applies morpheme internally (a claim Hayes (1986a:469) attributes to McCarthy). The OCP ensures that tautomorphemic geminates have the form in (3-7a), while the morphemic tier hypothesis (McCarthy 1989) ensures that heteromorphemic geminates have the form in (3-7b).

(3-7)



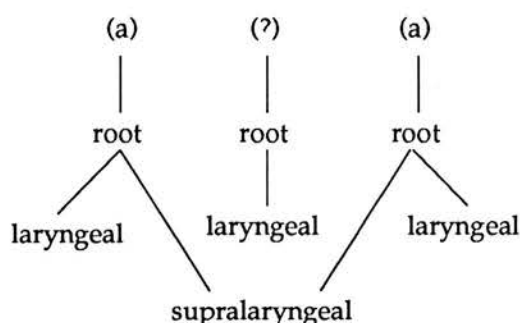
⁴ i.e. a geminate where the components come from different morphemes, such as the *nn* in *innumerable*.

⁵ i.e. a geminate where the components come from the one morpheme, such as the *nn* in *connect*.

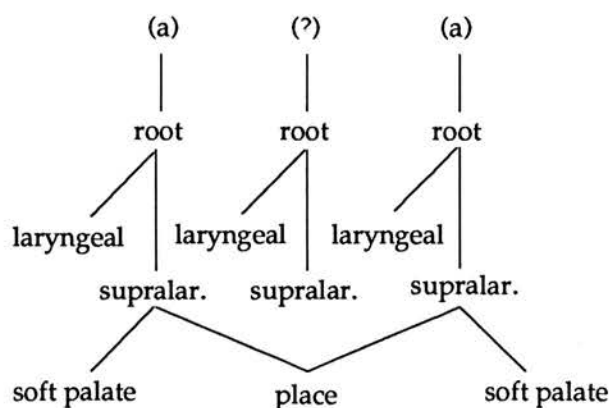
To analyse Icelandic preaspiration as the diphthongization of a true geminate therefore requires the additional assumption that heteromorphemic geminates have been modified from being of the ‘fake’ variety to being of the ‘true’ variety. Although such modifications are plausible, the assumption that such a rule exists should be accompanied by external supporting evidence, which is lacking here. In conclusion, existing analyses of the Icelandic data would seem to rely too much on controversial assumptions to be taken as clear evidence for the existence of the supralaryngeal (or the laryngeal) node.

Clements obtains further support for the existence of a supralaryngeal node from Klamath (Clements 1985:234–5). However, Iverson (1989:295ff) argues that a supralaryngeal node is not necessary for an analysis of Klamath (following an analysis given by Sagey 1987). Another example is from Acoma (Sagey 1986:34), where two vowels are normally identical when separated only by a glottal stop. Sagey argues that since the glottal stop lacks supralaryngeal features the vowels on either side can share their supralaryngeal specification.

(3–8) Acoma Analysis:



However, as for Klamath, another analysis is conceivable. Sagey’s supralaryngeal node dominates the feature nasal and a place node. In the absence of evidence about the sharing of nasality across a glottal stop, the Acoma data can equally well be analysed as the sharing of just the place node.



This observation accords with McCarthy's comment that "spreading of the supralaryngeal node, as distinct from spreading of the place node, is known from only one or two examples that are subject to plausible reanalysis" (1988:92).

As the above evidence is not particularly compelling I propose to follow McCarthy (1988) and Iverson (1989) in eliminating the supralaryngeal node. Picking up the theme of articulatory phonetic motivation once more, note that the supralaryngeal node does not correspond to any articulator or group of articulators. This lends support to the hypothesis that uncontroversial phonological evidence for the existence of a hierarchical node is correlated with the existence of a corresponding articulatory grouping.

The Manner Node. Although Clements acknowledges that "there is very little evidence to suggest that the manner tier functions as a unit" (1985:238) he still employs a hierarchical manner node and groups underneath it the features [continuant], [consonantal], [nasal], [sonorant], [strident] and [lateral]. Clements places the manner node with the place node underneath the supralaryngeal node. Given the absence of evidence for the manner node, Sagey (1986:45) assumes that these features do not form a constituent. However, Sagey only deals with the first three, leaving out sonorant, strident and lateral because, she argues, sonorant corresponds to a disjunction of properties, strident is an acoustic property, and lateral does not fit naturally into her conception of the place node (1986:280-1). Others have advocated different positions for these features. Here I shall simply review the claims and not the supporting

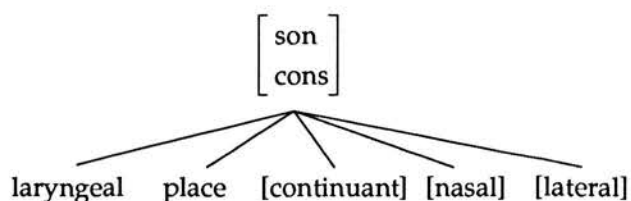
data. For readability the information is tabulated⁶.

(3–10)

Feature	Position	Reference
continuant	root	Sagey (1986), McCarthy (1988), Avery & Rice (1989), Paradis & Prunet (1989)
	supralaryngeal	Davis (1989)
consonantal	root	Sagey (1986), Paradis & Prunet (1989)
	inside root	McCarthy (1988)
nasal	root	Piggott (1987), McCarthy (1988), Paradis & Prunet (1989)
	supralaryngeal soft palate	Sagey (1986)
	supralaryngeal sonorant	Avery & Rice (1989)
	supralaryngeal	Hayes (1990)
sonorant	inside root	McCarthy (1988)
strident		
lateral	place coronal	McCarthy (1988)
	supralaryngeal sonorant	Avery & Rice (1989)

The proposal discussed by McCarthy (1988:97) differs significantly from the others in that it permits features to be *inside* a node. The features sonorant and consonantal never spread, delink or exhibit OCP effects unless they do so in concert with all the other features. However, if [sonorant] and [consonantal] are represented as normal constituents then there is nothing (principled) which prevents them from manifesting this unobserved behaviour. McCarthy's solution is to employ a structure of the following variety:

(3–11) **Class nodes as feature bundles:**



Here a hierarchical node has internal features. Note that this view is familiar from syntax (e.g. Gazdar et al. 1985) and its application to phonology has been suggested by Vincent (1986:317).

⁶This is not intended to give a complete picture. In particular, all suggestions to locate all or some of these features under a manner node (e.g. Iverson 1989, Hayes 1990) are omitted, as are possibilities which have been assumed without supporting evidence.

How are we to interpret these conflicting claims represented by the above table? There are a number of possibilities. There seems to be a widely held assumption that there is a single, universal organization of features, and that, given time, the empirical evidence will mount up sufficiently in favour of one option over all the others. Those phenomena which are left without the most natural node would need to be reanalysed. (Remember that feature hierarchy does not exclude any analyses but only makes some analyses easier to state than others.) However, from the perspective of so-called universal grammar, there is no reason to assume that phonology is not parametrized. For example, some languages could be specified 'nasal:supralaryngeal' and others as 'nasal:root' to indicate variation in the locus of the feature nasal. Furthermore, one could explore the potential implication relationships between various parameters or advance arguments in favour of particular parameter settings being 'marked' or 'unmarked'. A third option is to formulate statements of the form: "other things being equal, languages prefer the F feature to be linked to the N node" or "a grammar is more highly valued to the extent it employs a feature hierarchy where the F feature is linked to the N node". Such statements are widespread in phonology⁷. However, none of these three stances is attractive as a potential phonological theory since none are falsifiable. A fourth option is based on the observation that existing proposals for manner features fail to capture their atypical interpretation. Representing [sonorant] as a binary feature in the same way as (say) [round] fails to express the entirely different natures of these two properties. This fourth option is developed further in the next section.

To conclude the discussion of manner features, first note that most of the controversy represented by the above table concerns the supralaryngeal node. Clearly by dispensing with this node there will be a number of phenomena requiring reanalysis beyond those discussed above. However, the arguments for siting a feature under the supralaryngeal node typically assume the prior existence of the node, and so it

⁷ For example, Schein & Steriade (1986:741): "All other things being equal, association is preferred over copying"; Hayes (1986b:323): "I assume that the obligatory contour principle is a statement of markedness, assigning greater value to the structure of 4a"; Goldsmith (1990:323): "If we think of well-formedness—or its opposite, ill-formedness—as a matter of degree, then the path that a representation takes ... may be conveniently thought of as a downhill path towards a 'local minimum' of ill-formedness".

is difficult to interpret them directly as arguments for the node. Second, note that the manner features are only seen to create problems because of the assumption that there is a unique (non-parametrized) universal feature hierarchy. This assumption is ultimately based on the physiological endowment which is common to humans. Given the extreme difficulties of determining the validity of general claims about neurological organization—where do we begin looking for the supralaryngeal node?—it is more plausible to start from the assumption that a universal feature hierarchy, if it exists, arises from the fact that we all use essentially the same vocal apparatus.

The Place Node. The place features include the following: [coronal], [anterior], [distributed], [high], [low], [back], [round] and [labial]. Although the manner features do not function as a unit, there is considerable evidence to suggest that the place features do. Clements (1985:235–6) cites the case of place assimilation of /t, d, n/ in English and provides the following data:

(3–12)

–[θ]	[t]– eighth	[d]– hundredth	[n]– tenth, enthuse
–[tʃ], –[dʒ]	each, cheer	edge, gem	inch, hinge, insure, enjoy
–[r]	tree	dream	enrol

Here we can observe that the place of articulation of the stop agrees with that of the following segment.

Sagey (1986:37) provides the data from Kpelle (Mande, Western Sudanic, Liberia) which is reproduced in (3–13). Note that [k̠p], [g̠b] and [m̠ŋ] denote labio-velar articulations⁸. (We shall not be concerned with the progressive voicing assimilation here.)

(3–13)

/N-polu/	[m̠bolu]	“my back”
/N-tia/	[n̠dia]	“my taboo”
/N-kOO/	[ŋ̠gOO]	“my foot”
/N-k̠piŋ/	[m̠ŋ̠g̠biŋ]	“myself”
/N-fela/	[m̠vela]	“my wages”
/N-sua/	[n̠jua]	“my nose”

⁸ The symbol η represents the velar nasal.

Sagey notes that the nasal segment assimilates in all and only the place features. For example, manner and laryngeal features are not spread onto the nasal, for if they were the nasal preceding an [f] would be a fricative and voiceless. Furthermore, the resulting nasals are syllabic and are not grouped with the following segment to form a complex prenasalized segment, derived by linking the nasal feature to the following segment. Therefore the only possible account is regressive homorganic assimilation.

Given the relatively uncontroversial nature of this category I shall not survey any further evidence, but move on to various proposals for the structure below the place node. Before doing so, however, I shall pause to note the connection between the place node and phonetic structure. Sagey (1986:40) states that the place node does not correspond to an articulator but instead has acoustic motivation: "the distortions produced by place features have to do with changing the shape of the resonator". However, the three articulators responsible for the place features (i.e. the lips, tongue tip and tongue body) are all oral, and so are ultimately tied to a common articulator—the jaw—which has occasionally been implicated in articulatory models of speech production (e.g. Lindblom & Sundberg 1971, Mermelstein, Maeda & Fujimura 1971, Mermelstein 1973, Rubin & Baer 1981).

Sub-Place Groupings. Clements' model of the place node has it directly dominating all of the place features. However, Sagey (1986) has observed (citing Halle 1983) that this approach is inadequate for the representation of double and triple (oral) articulations. The relevant passage from Halle (1983) is reproduced below:

Consonantal occlusions are thus produced by three distinct active articulators: the lower lip, the front part of the tongue, and the tongue body. Since the position of each of these three articulators is independent of the other two it should be possible to produce consonants with more than one occlusion. Since there are three active articulators and since a given articulator can be at exactly one point at a given time there should exist three types of consonants with double occlusion and a single type of consonant with triple occlusion. (Halle 1983:99)

Halle provides the table in (3–14)⁹. Although Halle found no example of a triply occluded consonant, Sagey (1984) has claimed that such a consonant exists in Kinyarwanda¹⁰, and so this has been included for completeness. (Other triple articulations can be found in Shona, according to Sagey 1986:72.)

(3–14)

labio-velar	[k̠p]	Yoruba	[ak̠pa]	“arm”
labio-coronal	[p̠t]	Margi	[p̠tel]	“chief”
corono-velar	[ɰ]	Zulu	[a a]	“climb”
labio-corono-velar	[t̠kw]	Kinyarwanda	[t̠kwaan̠ga]	“we hate”

Sagey shows that previous accounts of sub-place structure (such as Clements’) are unable to express the fact that these are the only kinds of double or triple articulations that are possible. For example, a labio-coronal must be simultaneously [–coronal] and [+coronal]. Sagey provides a way to escape from this problem:

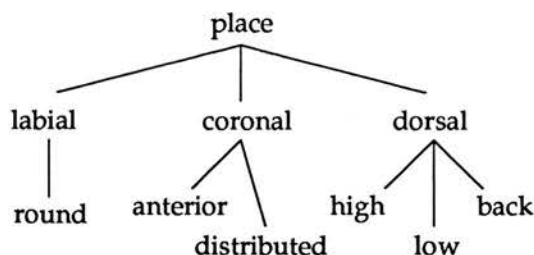
The solution to this problem lies in realizing that it is really irrelevant to the articulation of the labial closure (i.e. to the behaviour of the lips) whether or not there is additional [+coronal] closure. Therefore, a lack of coronal closure should not be part of the universal definition for a labial, indeed its defining characteristic, as it is when we define a labial as [+anterior, –coronal]. ... In short, the problem with the feature specifications in [Clements’ model] is that they define segments, not simply in terms of what constrictions or articulators are involved, but also in terms of what is not involved. (Sagey 1986:64–5)

Her solution has two parts. First, sub-place nodes are created for each of the three articulators: the nodes ‘labial’, ‘coronal’ and ‘dorsal’. Second, the applicability (or appropriateness) of the various place features is restricted to a particular articulator (see (3–15)). There are obvious phonetic justifications for doing this, but there are also phonological justifications, as surveyed by McCarthy (1988:103). For example, anterior is only specified for coronal segments and not for (say) labial segments as in previous models. The structure which results can be depicted as follows:

⁹ The ‘|’ symbol denotes the dental click, IPA: ‘turned t’. It is employed in English and sometimes transcribed *tsk*. Although clicks are occasionally considered to be ‘exoticisms’ (a term coined by Schane 1973:32) they are just as important as any other class of sounds or articulations for determining the feature hierarchy.

¹⁰ However, this might just be a cluster of two double articulations, namely t̠k and w, or t and kw.

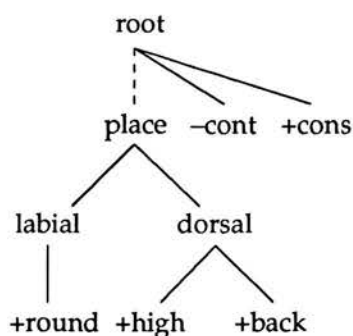
(3-15) Sagey's Place Node:



Segments having one (distinctive) place of articulation are only specified for one of these articulator nodes; a doubly articulated segment (e.g. $\hat{k}p$, w , u) is specified for two of these nodes, and so on. In other words, the presence of a labial, coronal or dorsal node denotes the active involvement of the corresponding articulator (Sagey 1986:67). For the many languages which only have singly articulated segments we merely need to stipulate that the place node only ever dominates one articulator node (at least at an underlying level).

Recall that manner features must be represented independently from place features. When non-distinctive (or predictable) articulation details are specified in the structure then all three articulator nodes are present and the information about which articulator the manner features concern is lost. For example, English coronal and velar consonants are often labialized when followed by a labial vowel (e.g. u) or glide (e.g. w). Compare the $[k]$ in *quit* with that in *kit*. In Sagey's hierarchy, the labialized $[k]$ might be represented as in (3-16) after a rule supplies the labial specification (omitting the laryngeal and supralaryngeal nodes).

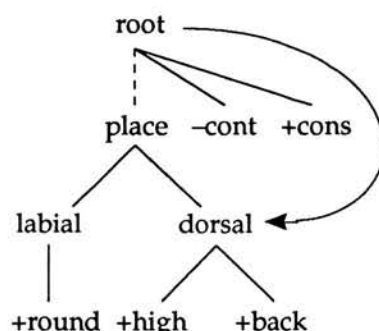
(3-16)



The information that the closure is dorsal and not labial is not represented in (3-16). Sagey's solution is to draw an arrow from the root node to the articulator node that

the manner features describe. Such an articulator node is called a major articulator (1986:203). The above example could be represented as in (3–17).

(3–17)



Double articulations have two major articulators and so have two of these arrows. This represents the fact that both articulators have the same degree of closure (1986:217). How Sagey intends this to work for the Kinyarwanda [tkw] is unclear, since the degree of closure for w is not the same as that for t or k.

Further arguments for Sagey's model have been advanced by Yip (1989b), who has surveyed the co-occurrence restrictions in homorganic consonant clusters in several languages. Problems remain, however, with the dorsal node. McCarthy has observed "the feature [dorsal] alone is obviously inadequate to characterize the degrees of freedom of the tongue body, and in particular it is an entirely unsatisfactory account of the interactions and lack of them between vowels and consonants" (McCarthy 1988:102). Currently there exists no consensus on the resolution of this problem.

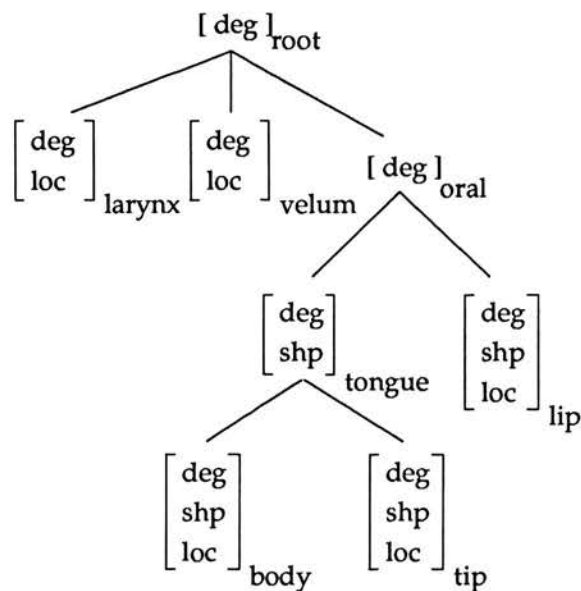
3.2 An Articulatory Model

It has already been claimed above that the structure of the vocal tract provides support for particular feature groupings. However, Clements (1985:230) states that "the ultimate justification for a model of phonological features must be drawn from the study of phonological and phonetic processes, and not from *a priori* considerations of vocal tract anatomy or the like". McCarthy (1988:89) makes a similar statement. For Clements and McCarthy, the fact that certain groupings are physiologically justified while others are not can only be an irrelevant artefact. However, there needs to be

a good explanation of why such feature hierarchies have so much in common with vocal tract organization while remaining distinct. Faced with this situation there are at least two responses: (i) simply ignore the physiological evidence or (ii) heed the evidence and attempt to incorporate the areas of mismatch in some other way. In this section the latter option will be explored in detail, following the proposals of Browman & Goldstein (1989).

The hierarchical organization I shall assume is depicted in (3–18). Each node consists of a feature bundle¹¹, where [deg] is a (polyvalent) feature for constriction degree, [loc] is a (polyvalent) feature for constriction location and [shp] is a (binary) feature for articulator shape (lip rounding or tongue laterality). Each bundle has a sort, specified as a subscript (more on sorts later). The motivation for the hierarchical organization is articulatory. Each node corresponds directly to an articulator (except for the oral node)¹².

(3–18)



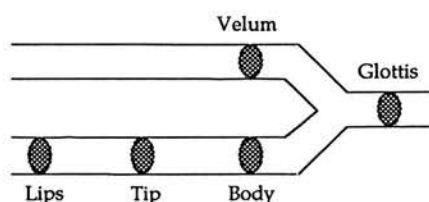
The [deg] feature corresponds roughly to the former manner features, except that it

¹¹ Recall from chapter two section one that sorts are classifications of nodes. In general, there is no limit to the number of sorts which can be true of a node. When a node has more than one sort, it will often be convenient to depict its sorts using feature bundles.

¹² It is hypothesised that recognizing a location parameter for laryngeal constriction will permit ingresses and ejectives to be described. As the raising or lowering of the larynx does not play a significant role in the later discussion this feature will usually be omitted. For completeness, the velum is also given a constriction location, but since this location is invariant the specification will always be omitted.

is present on *every* node. The value of this feature is percolated through the tree: its value at a node is constrained by its value at the daughters and the mother of the node. The percolation of constriction degree is best understood in terms of 'tube geometry' (Browman & Goldstein 1989:234ff). Imagine the vocal tract as a collection of interconnected tubes along with a number of 'valves'. This can be represented schematically as in (3-19).

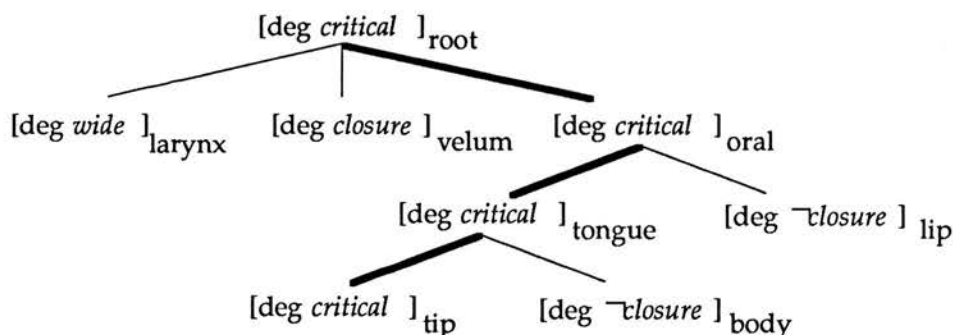
(3-19)



Each valve will have a number of settings. When a valve is closed, we shall say that the corresponding section of tubing is maximally constricted. This setting will be called 'closure'. The other settings (in order of decreasing constriction) are 'critical', 'narrow', 'mid' and 'wide', following Browman & Goldstein (1989). The last of these corresponds to minimal constriction. The constriction degree of a pair of tubes connected in parallel is the minimum of the individual constriction degrees of the two tubes. If the tubes are connected in series then the overall constriction is the maximum of the individual constrictions ¹³.

To illustrate this percolation, consider the segment [s]. A representation appears in (3-20), where the thick lines record percolation of degree specifications.

(3-20)



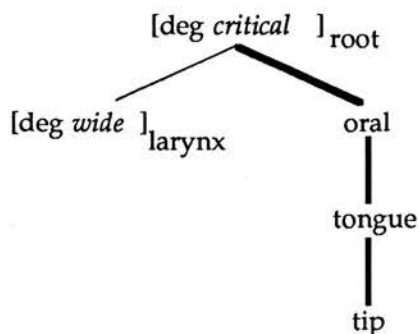
¹³This use of maximum and minimum differs from that of Browman & Goldstein (1989:237), but it is adopted as it seems more natural for a stop to correspond to maximal constriction rather than minimal constriction.

This segment involves a tongue tip constriction with the value *critical*. So long as there is no tongue body closure, the effective constriction degree of the whole tongue is also *critical*. A similar situation holds for the oral node. For the segment to be voiceless and non-nasal the constriction degree of the larynx must be *wide* and the velum constriction degree must be *closure*. As we shall see later, the effective constriction degree of the entire vocal tract is computed as follows:

(3-21) root:deg
 = max(larynx:deg, min(velum:deg, oral:deg))
 = max(wide, min(closure, critical))
 = max(wide, critical)
 = critical

Note that negative information ([deg -closure]) was specified in the above tree. This required that, whatever the lip and tongue body constriction degrees are, neither is closure. This negative information was specified so that the percolation of constriction degree could be illustrated. (As we shall see later, the underlying representation of [s] will not normally contain this negative information.) The precise details of underlying representations depend heavily on the phonology of the language in question. However, one possible underlying form for [s] is given in (3-22).

(3-22)



Again, the thick lines show the percolation. Here we view the thick lines as actually part of the structure, specifying (rather than illustrating) which path percolation must follow. Therefore it is only necessary to represent the constriction degree in one node, and the value can percolate up or down from that node¹⁴. In (3-22) it is specified

¹⁴ The percolation metaphor is perhaps somewhat misleading here. Percolation is normally assumed to be in an upwards direction. When used in phonology, however, the term does not seem to carry any particular directional significance. For some, (e.g. Hayes 1990:44), percolation is only downwards. My own view of feature 'percolation' (described later) is equational, and equations are not directional.

in the root, and the oral, tongue, lip, body and tip constriction degrees of (3–20) are produced automatically. More details of this control of percolation will be given below. For now I shall just observe the connection between the path of emboldened lines in (3–22) and Sagey’s arrow notation, discussed in the previous section.

In the above example we saw that the [deg] feature is polyvalent. The [loc] feature is also polyvalent. In the next table the values of each are provided. The unmarked (or default) value of each is shown in boldface.

(3–23)

Node sort	Constriction Locations	Constriction Degrees
larynx	raised, lowered	wide, critical , closure
velum	velic	wide, closure
body	palatal, velar , uvular, pharyngeal	wide , mid, narrow, critical, closure
tip	labial, dental, alveolar , post-alveolar, palatal	wide , mid, narrow, critical, closure
lips	protruded, labial , dental	wide , mid, narrow, critical, closure

The tongue, oral and root nodes can also have a constriction degree. I shall assume that these are the same as for the tongue body. These location and degree possibilities follow Browman & Goldstein (1989), except for those in the larynx. Here I am assuming that the raised or lowered state of the larynx is independent of the state of the glottis. Only three glottal states are provided, following Ladefoged (1971:17) who, while distinguishing nine points on the continuum of glottal constrictions, claims that “no language makes more than three oppositions on this continuum”. I take this claim to mean that no language requires more than three glottal states to be distinguished (“three oppositions” means 2^3 , or 8 states). The separation of lip rounding from lip constriction degree is based on the observation that vocalic lip rounding can begin well before the other components of vowel articulation (Bell-Berti & Harris 1979). Furthermore, Mermelstein (1973:1073) notes that these parameters have different characteristic rates of change (see also Browman & Goldstein 1986).

Browman & Goldstein (1989:240) have shown how representing constriction degree at every level permits an elegant account of natural classes. Translating their suggestion into the present framework, the following traditional manner features can be

expressed: [sonorant], [continuant], [consonantal] and [nasal]. In the following table, the constriction degree ‘open’ covers for *narrow*, *mid* and *wide*; ‘not open’ corresponds to *closure* and *critical*¹⁵.

(3–24)

Structural Encoding	Manner Feature	Natural Class
root:deg = open	+sonorant	sonorants (nasals, liquids, glides, vowels)
root:deg = not open	–sonorant	obstruents (stops, affricates, fricatives)
oral:deg = closure	–continuant	nasals, stops, affricates
oral:deg = not closure	+continuant	liquids, glides, vowels, fricatives
oral:deg = open	–consonantal	vowels, glides
oral:deg = not open	+consonantal	obstruents, liquids, nasals
velum:deg = open	+nasal	nasals
velum:deg = closure	–nasal	orals

The only manner features of Clements that this leaves out are [strident] and [lateral]. The feature [strident] distinguishes dental fricatives and affricates from alveolar ones; here these are distinguished using the [loc] feature. The feature [lateral] is represented using our (binary) feature [shp] in the tongue, tip and body nodes¹⁶.

Now we are in a position to further explore the percolation of constriction degree and its relation to the emboldened line notation. The latter notation will be used to express a ‘head-of’ relation¹⁷. A node which is connected by an emboldened line to its parent node will be considered the head constituent of that particular local tree. Nodes not so connected are not head constituents. A head node will, by definition, have the same constriction degree as its parent. Furthermore, this must be compatible with the definition of percolation. For example, if all the constituents of a node are connected in parallel (cf. 3–19), then the degree of the head constituent must be at

¹⁵ The cover features ‘open’ and ‘not open’ are formalized in section four.

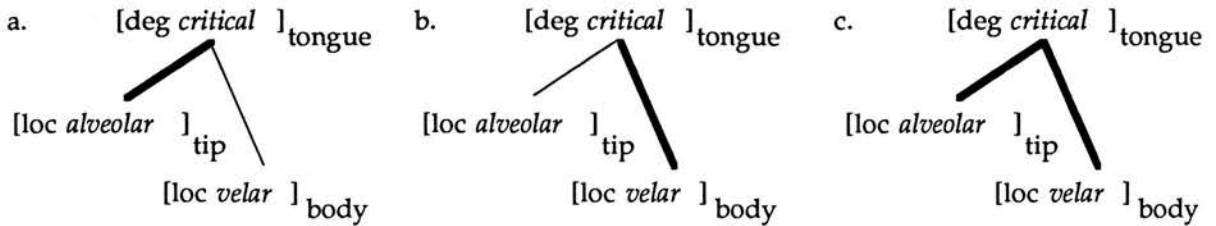
¹⁶ By requiring that any specification of [shp] on the tip and body nodes must agree with the specification of [shp] on their mother node (the tongue node), an agreement effect is achieved. This is based on the assumption that laterality is a property of the whole tongue and the hypothesis that no languages manifest two independent laterality contrasts, one coronal and one dorsal. Note that this hypothesis differs from the claim that a lateral tongue tip gesture cannot be coarticulated with a non-lateral tongue body gesture. An apparent counter-example to the latter hypothesis is provided by the English [kl] cluster, which may involve some overlap between the two tongue gestures. However, this raises the question of the interpretation of laterality versus centrality when there is complete closure. I do not address this matter here.

¹⁷ The use of the term here is only for convenience, and does not embody the claim that such notions as ‘head’ and ‘projection’ from syntactic theories should be imported into phonology. In particular, my use of headedness concerns unpredictable rather than predictable information.

least as constricted as the degrees of all the other constituents.

Consider the case of the tongue node. Suppose that it has a specified constriction degree of *critical* (owing to percolation). There are three possibilities to consider, exemplified in (3–25).

(3–25)



In (3–25a), the tongue tip is the head constituent and therefore has a critical constriction degree. Since this must be compatible with percolation, the tongue body constriction degree can be anything other than *closure*. A similar situation holds for (3–25b). In (3–25c) both tip and body nodes are *critical*. (This may be a suitable representation for a doubly articulated voiceless fricative, as is said to occur in Swedish for orthographic *kj*, Pullum & Ladusaw (1986:69)). Recall that Sagey’s notation permits more than one arrow, in order to represent segments with two major articulators, and the manner features applied equally to both articulators.

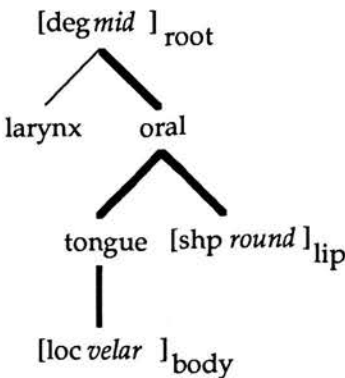
The situation for the oral node and its two constituents tongue and lip is analogous to the above, and need not detain us further. The final case we must consider is that of the root node, which dominates larynx, velum and oral nodes. Recall from Browman & Goldstein’s (1989) ‘tube geometry’ (discussed in the previous section), that computing the effective constriction degree of the vocal tract is a two stage process. First the effective constriction degree of the supralaryngeal section of the tract is computed as the minimum of the constriction degrees of the oral and velum nodes. Next, the maximum of the supralaryngeal and the laryngeal constriction degrees is computed. This is expressed informally as in (3–26), a format we have already seen.

(3–26) $root:deg = \max(larynx:deg, \min(velum:deg, oral:deg))$

Examples.

Consider the case of English back vowels. The way in which the degree of rounding and the vowel height are correlated is captured in the following representation. (Clearly the jaw plays some part in this observation, and the present approach fails to express that adequately.) In (3–27) the constriction degree is *mid* (corresponding to the IPA cardinal vowels 6 or 7).

(3–27)



For a higher vowel (such as [u]), the overall constriction degree will be *narrow*, and this will be percolated to both the lips and the tongue body.

The next example is intended to illustrate the use of the tongue node in an analysis. The doubly articulated (egressive) stops of Igbo (and presumably many other languages as well) are all linguo-labial. Igbo does not have stops like [k̂t], which are corono-velar. This observation can be stated (informally) in the current framework as follows:

(3–28) The tongue node has at most one head constituent.

It would seem that this kind of constraint can only be neatly stated in a hierarchical organization that employs a tongue node. It may even be that languages only have [k̂t] if they also have [p̂t] and [k̂p]. This observation, if true, could be stated as an implicational universal thus:

(3–29) If a tongue node can have more than one head constituent, then so can the oral node.

Tying constriction location and degree together at the level of the leaves in these tree structures would appear to go against the abundantly exemplified claim that

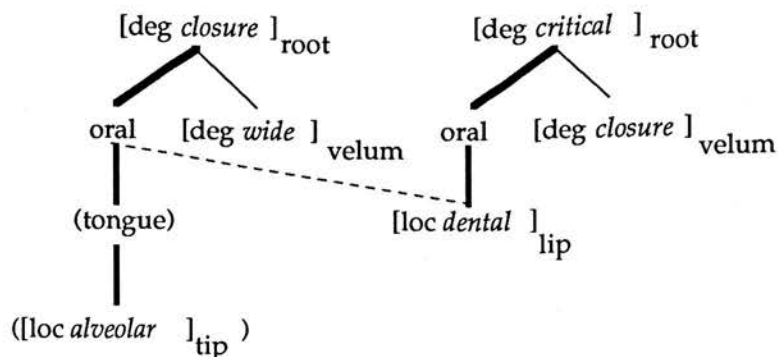
place and manner features are independent. Consequently, it may seem that the organization advocated here will not be able to deal appropriately with place assimilation phenomena. There are two ways around this problem. One is to postulate that in certain cases where place features are thought to spread independently of manner features, they actually spread together, but that the spread manner features have a relatively minor acoustic effect. Consider the data Sagey (1986:37) presented from Kpelle, already discussed in the above section on the place node. The data is reproduced below:

(3-30)	/N-polu/	[ɪnbolu]	"my back"
	/N-tia/	[ɪndia]	"my taboo"
	/N-kOO/	[ɪŋgOO]	"my foot"
	/N-kpiŋ/	[ɪmŋgbiŋ]	"myself"
	/N-fela/	[ɪnvela]	"my wages"
	/N-sua/	[ɪnjua]	"my nose"

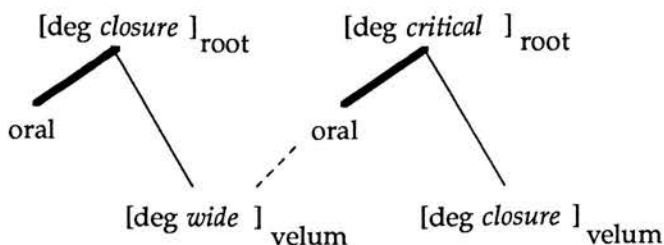
Sagey observes that "only place features, and not manner or laryngeal features, are spread onto the nasal. /f/ conditions a labial nasal stop, not a labial nasal fricative". Presumably this labial nasal stop is actually labio-dental. However, Catford claims that the labio-dental nasal "is probably realized most frequently as a nasalized approximant rather than the usual type of nasal, which requires an airtight oral closure" (Catford 1988:85). From an aerodynamic point of view, a nasalized fricative clearly requires a relatively narrow velic aperture. If the velic aperture is wide then almost all of the air will be released through the nasal passage. The amount escaping through the narrow oral aperture will not be sufficient to produce turbulence (Poser, pers. comm.). Browman & Goldstein (1989:242) make essentially the same observation, showing how it follows from constriction degree percolation. Given these phonetic facts, Sagey's observation that /f/ conditions a labial nasal stop and not a labial nasal fricative does not justify her conclusion that the manner features must not spread.

At least two analyses of the Kpelle data are conceivable in the present framework. The first employs dominance and the second employs overlap. (I shall not try to choose between these two analyses here.)

(3-31) a. **Dominance solution:**



b. **Overlap solution:**



The dotted line in (3-31a) indicates the creation of a dominance line between an oral node and a lip gesture. The dotted line in (3-31b) indicates the overlap of the nasal gesture of the first segment with the oral gesture of the second. The oral closure for the first segment is assumed to be unspecified at this level. However, as Browman & Goldstein note, the oral closure of this segment could be completely hidden by that of the following segment (see chapter 1 section 3 for some discussion of acoustic hiding).

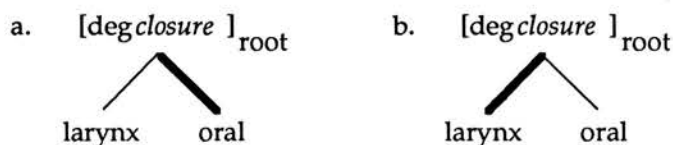
The importance of the above example is that the independence of place features from manner features may not be as securely established as it might have been thought, bearing in mind the heavy reliance of arguments for this independence upon cases of nasal assimilation. Of course, there still remain many cases of place assimilation which do not involve nasals (e.g. the English coronal assimilation data discussed above). Furthermore, some deletion and OCP phenomena would seem to require place/manner independence. A possible response to the former is discussed below, but the latter is not addressed here¹⁸.

¹⁸ There are a number of OCP effects which would seem to be sensitive to place features independently of any others, such as for Arabic consonantal roots (McCarthy 1981). However, form and precise

The second possible response to the place/manner independence is to regard downwards percolation of constriction degree as a default mechanism. Consider the case of voiceless consonant reduction to glottal stop in English (discussed above). Lass (1976:155) treats this as a two stage process, namely the deletion of the oral submatrix, along with the transfer of a [-cont] specification from the oral submatrix to the laryngeal submatrix.

My analysis of this consonant reduction is based on the alternation between (3-32a) and (3-32b). The generalization (here, disjunction) of these two will be taken to be the (partial) underlying representation of English voiceless stops. The structure in (3-32a) is intended to be the default, and that in (3-32b) will arise in various phonological environments (cf. Hudson's 1980 treatment of alternations).

(3-32)



I shall assume that the laryngeal constriction degree is filled in as *wide* by default¹⁹.

This concludes the discussion of (subsegmental) hierarchical feature structure. In this section I have attempted to marry Sagey's proposals with those of Browman & Goldstein. In the next section I shall suggest a way of linking up these subsegmental structures with suprasegmental ones, while at the same time doing away with the segment as a distinct level of hierarchical structure.

3.3 Prosodic Structure

In the first section of this chapter a survey of proposals for hierarchical feature organization was presented. The underlying assumption in all of the work reported there is that the root node corresponds to a segment, and that the hierarchical structure pertains to properties of segments. These properties (or features) have a considerable

function of the OCP is still a controversial matter, and it has not been addressed in the present study.

¹⁹ Note that this alternation of stops with the glottal stops only occurs for voiceless stops. An adequate account of this data must give some explanation for the non-involvement of voiced stops.

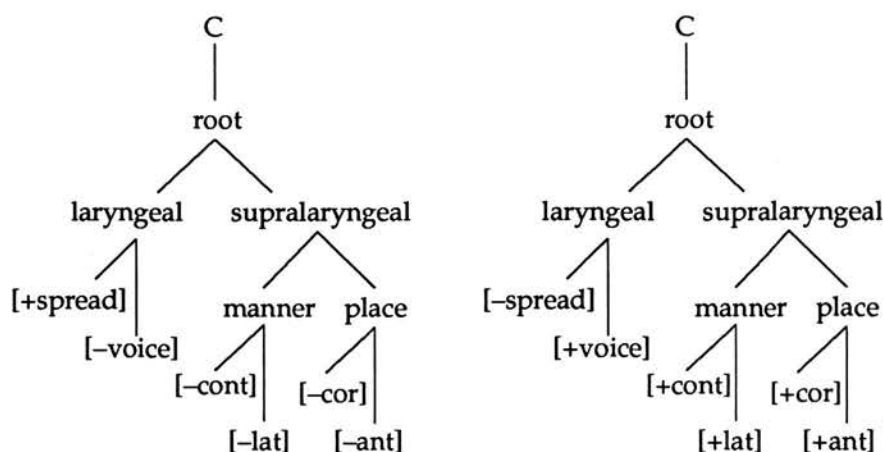
degree of autonomy, in that phonological rules can be sensitive to, and can manipulate, the features and feature groupings independently. However, the overall picture remains segmental. The discussion in the previous section, while continuing in the same vein, was not so clearly tied to the segment. In this section I shall endeavour to rid the segment from the proposed theory.

The present discussion will be motivated by facts from English phonotactics. Fudge (1987) lists all the possible English syllable onsets. However, he does not distinguish word-initial onsets from word-internal onsets. (This omission would seem equivalent to the naïve view of the collection of word junctures of an utterance being a subset of the collection of syllable junctures.) Word-internal onsets are a more reliable guide to syllable structure—assuming there is some independent way of locating syllable junctures—because word-boundary phenomena (such as ‘extrametricality’) complicate the picture as far as word-initial onsets are concerned. In general, the fricative-stop and fricative-nasal clusters do not appear in word-internal onsets except in compound words. This leaves the following onset clusters.

- (3-33) a. pr, tr, kr, br, dr, gr
b. pl, kl, bl, gl
c. tw, kw, dw, gw

We can make a number of observations about these clusters. All have two elements, where the first is a stop and the second is a liquid or a glide. Clusters involving *l* or *w* involve two articulators. The clusters with *r* which employ only one articulator (i.e. *tr*, *dr*) also employ only one place of articulation. The only voicing distinction manifested by the clusters is borne by the stops, i.e. voicing is a prosody of the clusters. Often overlooked is the fact that the ordering of consonants in these clusters (and more generally in word-initial clusters) is completely predictable.

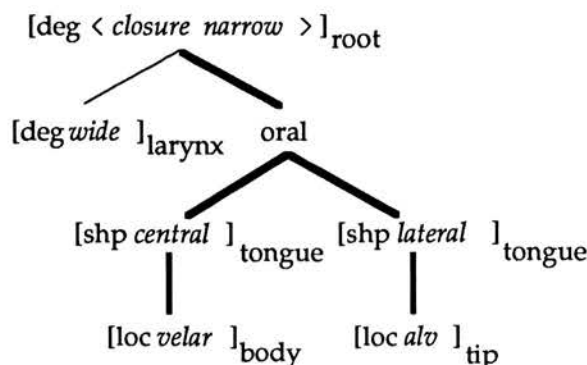
A segment-based analysis of this data would have the following appearance, here exemplified using Clements’ (1985) categories for the cluster *kl*:



Here the ordering (i.e. *kl* versus *lk*) is redundantly specified (in the context of an overarching syllable structure). The fact that *tl*, *dl*, *pw* and *bw* do not occur would probably need to be expressed as an OCP effect, or as the result of a dissimilation process. (Clements' (1985:240–1) discussion gives an indication of how this might be done.) Sagey's approach would also involve the use of two root nodes for these clusters, as they cannot be analysed as double articulations; the component articulations need different manner specifications.

A different view is afforded by the new perspective. Suppose that a root node can bear a sequence of constriction degree specifications, and that where a node dominates two head constituents of the same kind (and therefore sequenced by \prec) the first constriction degree is percolated to the left constituent and the second is percolated to the right constituent. Below is a possible representation for the *kl* cluster²⁰.

²⁰ Note that it would not have been possible for (3–35) to be formulated with branching at the tongue node instead of at the oral node. Branching at the tongue node would be to nodes of different kinds, namely a body node and a tip node, for which our percolation strategy will fail, eliminating this possibility. (Recall that when two nodes of different kinds are both head constituents then they form part of a complex segment, having the same constriction degree for both articulators.) There is nothing to prevent branching at the root node. Both constituents would be of the same kind, namely *oral* and so percolation would function correctly. However, branching at this relatively high level will be employed in the representation of consonant clusters, as we shall see later.



The root node contains the constriction degree sequence $\langle \text{closure narrow} \rangle$. I assume that this will be lexically specified as a set, and that its ordering will arise from a general sonority constraint. Notice that there is just one larynx specification for the cluster. Although the *l* (and also the *r* and *w*) may be at least partially voiceless, there is no voicing distinction for these ‘segments’, and it therefore seems reasonable to assume that the alignment of the glottal widening will be phonetically determined.

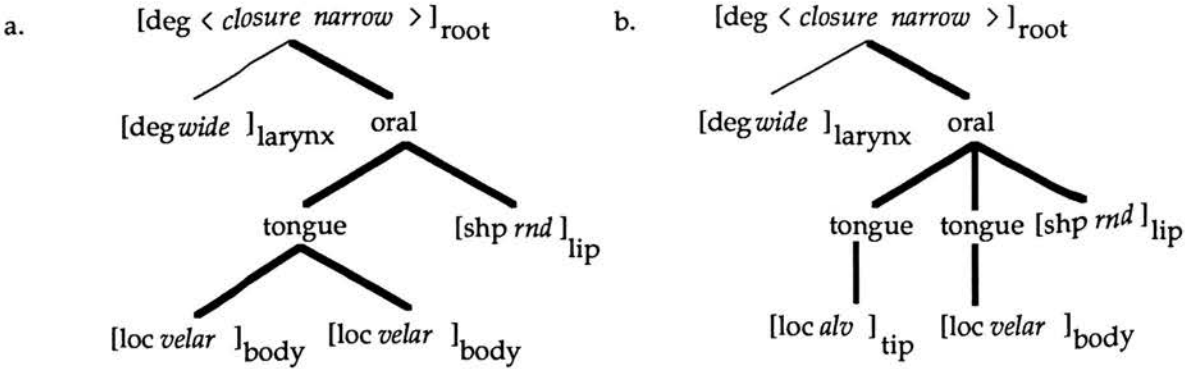
Now we have an account for the non-existence of *tl* and *dl* onset clusters (and possibly also *lt* and *ld* (word-internal) coda clusters). Since these both involve the tongue tip they must be dominated by the same tongue node. However, the tongue node dominating *d* must be distinctively specified [shp *central*] whereas the tongue node dominating *l* must be distinctively specified [shp *lateral*]. As these nodes have an incompatible specification for the shape feature they cannot be identical. A similar argument can be applied in the case of **pw* and **bw* if we assume that *p* and *b* are specified as nonrounded (for English).

Also necessary for English onset and coda clusters is the constraint that they must not contain complex segments such as $\hat{p}\hat{t}$ or $\hat{k}\hat{p}$, for example. However, since English back round vowels have been analysed as double articulations, this constraint only applies to stop consonants. For this result it is necessary to add the constraint that any *closure* constriction degree specified in the root node is only percolated to one articulator. In other words, the percolation of *closure* at each node can only be to a single constituent²¹. With this additional constraint we can provide structures for *kw*

²¹ A similar constraint was discussed in connection with the doubly articulated stops of Igbo above, where the tongue node could only contain one head constituent, not two (which would be required for

and *tw*, as in (3–36).

(3–36)



The constriction degree of the root node is a sequence in both (3–36a) and (3–36b), and both oral nodes inherit this sequence. How are the constriction degrees to be passed further down each tree? Consider the case of (3–36a). We have already seen that if a node dominates two nodes of the same type, these two nodes must be ordered, and a constriction degree sequence having two values will be split in two. When a node dominates two nodes of different sorts (as for the oral node in (3–36a)), both nodes receive the constriction degree of the parent node. This was the case when the parent node had only a single specification for constriction degree (e.g. as for the English vowels discussed above). However, this second kind of percolation may conflict with the above requirement that the closure specification may only be percolated to a single constituent. Therefore, the second kind of percolation needs to be modified to encompass both situations. It is presented in (3–37), along with a statement of the simpler, first kind of percolation, for the sake of completeness²².

a coronal-velar stop).

²² Although this definition is stated just for English, one may wish to employ it for other languages as well.

(3–37) **Percolation:**

- a. If a node P dominates two nodes Q and R of the same sort, then the constriction degree of P must be a two-element sequence S. The first element of S is equal to the constriction degree of the left constituent (say Q) and the second equal to the constriction degree of the right constituent (say R).
- b. If a node P dominates two nodes Q and R of different sorts, then the constriction degree sequences of Q and R must both be non-empty subsequences of the constriction degree sequence of P. Moreover, the constriction degree sequences of Q and R must exhaust the constriction degree sequence of P, and the intersection of the constriction degree sequences of Q and R must not contain a *closure* specification.

With this revision we can now consider the percolation possibilities for the constriction degrees of the oral node in (3–36). The *closure* constriction degree can only be passed to one of the constituents. If this is the lip constituent then the tongue constituent has only a single constriction degree. However, the tongue node must have two constriction degrees as it dominates two nodes of the same sort (namely, *body*). Therefore, the *closure* constriction degree must be percolated from the oral node to the tongue node. As the tongue node requires two constriction degrees, the *narrow* constriction degree must also be percolated down from the oral node. The lip node must have a constriction degree specification and it cannot be *closure*, so there is only one other option, namely *narrow*. Now the tongue node has a sequence of constriction degrees. The first of these (*closure*) is passed down to the left constituent, and the second (*narrow*) is passed down to the right constituent. The resulting structure has a velar tongue closure followed by a velar opening (to a narrow width); both of these articulations are simultaneous with a narrow rounded lip aperture, as required for *kw*.

Now consider the case of (3–36b) for *tw*. Again, the oral node has a sequence of two constriction degree specifications. The first of these (*closure*) must be percolated down to the left tongue node, and the second (*narrow*) to the right tongue node. What constriction degree does the lip node receive this time? By definition, it must receive at least one constriction degree, and it cannot be *closure* as *closure* must be present on another node, and *closure* can only be present on one node (for English). Therefore the lip node receives the *narrow* specification. The resulting structure has

an alveolar tongue closure followed by an alveolar opening (to a narrow width); both of these articulations are simultaneous with a narrow rounded lip aperture, as required for *tw*.

This concludes the discussion of English consonant clusters. It must be stressed that this discussion is primarily intended to illustrate various percolation options, and to show that it may not be necessary to have more than one root node for a consonant cluster. The details of percolation, while seemingly plausible, must be tried on a wide range of data before they can be accepted with any confidence.

Affricates. A widespread view of affricates is that they are contour segments, containing two manner specifications ([–cont] and [+cont]) for one place specification. This rests on the claim that affricates function as stops when ‘viewed from the left’ but as fricatives when ‘viewed from the right’ (Sagey 1986:29). However, this view is difficult to maintain in the light of the evidence (see, for example, Goldsmith 1990:68–73). Goldsmith claims that it is generally not the case that affricates manifest the dual function described above. Furthermore, some languages (e.g. most dialects of Spanish) have affricates with no corresponding fricatives. Therefore, analysing a Spanish affricate as a stop-fricative sequence introduces a segment (i.e. a fricative) which cannot appear on its own.

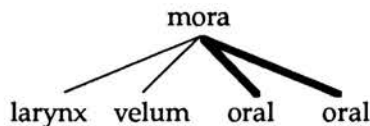
As it happens, the view of affricates as contour segments cannot be adequately encompassed within the proposal as it currently stands. It is not possible for a single leaf node (such as a body, tip or lip node) to contain two constriction degree specifications. However, it would be possible to have two leaf nodes of the same sort where one has a *closure* constriction degree and the other has a *critical* constriction degree, but then the place of articulation would be (redundantly) specified twice. Instead, I propose to view affricates as stops, and to distinguish them from other stops using place of articulation (loc) and shape of articulator (shp). For example, [tʃ] (or č) is a palato-alveolar stop, as distinct from [ʃ], a palato-alveolar fricative, and [t], an alveolar stop. [ts] and [t] contrast in apicality (represented using shp). This would seem adequate for Caucasian languages like Avar and Abkhaz, which have both an alveolar series and a palato-alveolar series, with stops, fricatives and

affricates, where the affricates are grouped with the stops (Lass 1984:152–3).

Moras and Syllables. In chapter one, section three, I outlined a moraic theory of syllable structure where the onset consonant(s) are grouped together with the first (perhaps only) vowel of a syllable into a mora. This mora is dominated by a syllable node, which may optionally dominate a second mora. The second mora, like the first, may contain consonants or vowels, or both. I am now going to equate this mora node with the root node. Since moras are dominated by syllables, the label ‘root’ will no longer be appropriate, and henceforth the label ‘mora’ will be used exclusively.

Vowels are typically voiced, and typically do not manifest a nasality distinction that is not also a part of the surrounding environment. Therefore, I shall assume that a mora dominates only one larynx node and only one nasal node, but that it can dominate more than one oral node, as illustrated in (3–38). We have already seen multiple branching to nodes of the same sort *below* the level of the oral node, but until now there has been no use for branching at this level²³.

(3–38)



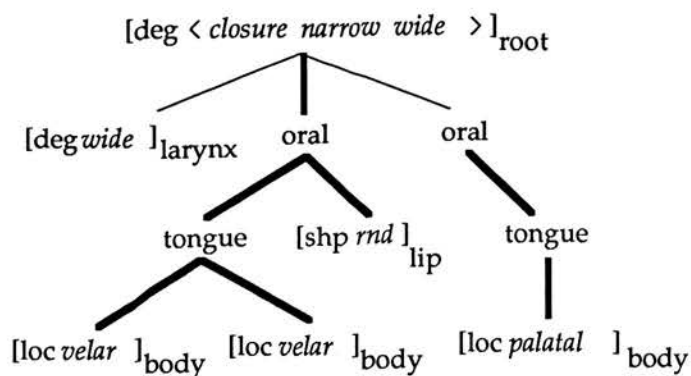
By the definition of percolation, such a structure must have two constriction degree specifications in its mora node, so that the first can be passed to the left oral node and the second to the right oral node. However, if an oral node can also contain branching to nodes of identical sorts (as required for consonant clusters and similarly for diphthongs), it will need to have a sequence of constriction degree specifications itself. This motivates a further change in the percolation strategy, which is stated in (3–39).

²³ Note that structures of this kind do not carry the ‘segmental’ implication that everything under the first oral node precedes everything under the second oral node. In fact, there could be a high degree of overlap between the constituents of the two nodes. Furthermore, there is no implication that the single larynx and velum specifications must last for the combined duration of both oral specifications. Rather, the claim is that there is only one distinctive laryngeal specification and one distinctive velum specification per mora. The temporal locus of this contrast is not a phonological matter but a phonetic one.

- (3-39) If a node P dominates two nodes Q and R of the same sort, then the constriction degree of P is a sequence $\langle c_1, \dots, c_n \rangle$ where n is greater than 1. The constriction degree of Q and R are $\langle c_1, \dots, c_m \rangle$ and $\langle c_{m+1}, \dots, c_n \rangle$ respectively.

In other words, the constriction degree sequence of a node is *shared out* to all its (head) daughter constituents when these latter are of the same kind. This richer notation of percolation may need to be further refined to allow for n-ary branching (rather than just binary branching). However, there is no obvious need for this as yet. As an example of the use of binary branching to oral nodes, consider the representation of /kwe/ (as it occurs in an English word like *quest*) depicted in (3-40). (Note that we have already seen the /kw/ part of this in a previous example.)

(3-40)



The three constriction degrees of the mora node must be shared out between the two oral nodes. There are two possibilities. Either the first oral node has two constriction degrees and the second oral node only has one, or vice versa. Clearly only the first of these possibilities will work, given that the tongue node must have two constriction degrees. Notice again that the constriction degree specifications of the oral node progress in the order of increasing sonority. If we are given that this mora is the first mora of a syllable then this increasing sonority is predictable and need not be stated explicitly.

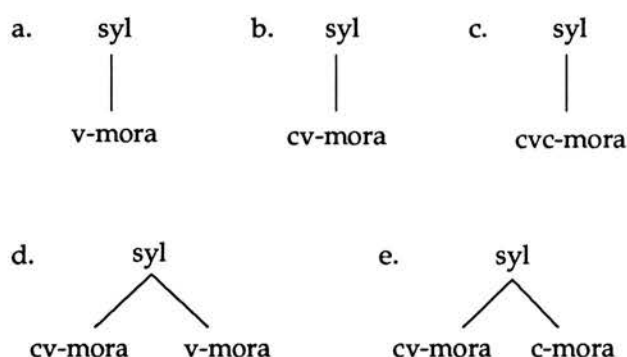
There are a number of possible sequences of constriction degrees that a mora can have. However, for the purpose of syllabifying moras, less detail will be required, as we shall see in due course. I propose the following classification of constriction degree sequences:

(3-41)

- c-mora(x) Each element of the sequence x is *closure* or *critical*, and each element is equal to or narrower than the last.
- v-mora(x) Each element of x is *narrow* or *wider*.
- cv-mora(x) Each element of x is wider than the last. The first must be *closure* or *critical*, and the last must be *narrow* or *wider*.
- cvc-mora(x) The sequence x can be divided into two sequences y and z such that cv-mora(y) and c-mora(z) are both true.

These are classifications of sequences of constriction degree specifications. However, I shall instead view them as classifications of moras, depending upon their constriction degrees. Now these classifications can be used in the development of constraints on syllable structures. The next example gives a characterization of a range of different syllable types.

(3-42)



The syllables in (3-42a,b,c) are called *light*, as they contain only one mora, whereas those in (3-42d,e) are called *heavy*, and contain two moras. The syllables in (3-42a,b,d) are called *open*, as their rightmost constituent is a vowel, whereas those in (3-42c,e) are called *closed*, and have a consonant as their rightmost constituent. Just as we used the sorts *c-mora*, *v-mora*, *cv-mora* and *cvc-mora* to classify moras, we can use the sorts *light*, *heavy*, *open* and *closed* to classify syllables. This general approach could be applied to the higher levels of prosodic structure, constraining the number and variety of syllables in metrical feet, and so on. One further definition is necessary before we can go on to discuss licensing. The first mora in a syllable will be given the sort *onset*. The second mora, if there is one, will have the sort *coda*²⁴.

²⁴ A caveat is necessary here. The light closed syllable in (3-42c) has an onset and a coda, but only

Autosegmental Licensing. Further constraints on the internal structure of moras come under the heading of AUTOSEGMENTAL LICENSING (Goldsmith 1990). The paradigmatic rôle of consonants often varies depending on the position of the mora in the syllable. In many of the world's languages, a wider variety of distinctions are made in the 'onset mora' than in the 'coda mora' (where there is one). Goldsmith (1990:125) claims that "any feature that is distinctive in the language can appear in at least one position in the onset of nucleus. If the latter were not so, there would be a feature that could appear distinctively only in the coda; but that never occurs."

One of the most common observations encompassed by autosegmental licensing is that the articulation location of coda consonants is non-distinctive in many languages (e.g. Hausa, Goldsmith 1990:128). Coda consonants do have an articulation location, but this is determined by the onset of the following syllable (or by a default if there is no following syllable). The statement that articulation location is licensed only by an onset mora is expressed using the following constraint.

- (3-43) A place of articulation specification must be the head of an onset mora. i.e. there must be a path from an onset node to the relevant articulator node which only involves **emboldened lines**.

This approach appears simpler than the one adopted by Itô (1989:224). Her analysis employs the statement in (3-44) along with Hayes' (1986b) 'linking constraint' which requires the association lines mentioned in structural descriptions to be exhaustive. Note that the precise interpretation of [PLACE] in (3-44) is unclear²⁵.

- (3-44) **Itô's Coda Filter:**



The force of the linking constraint is to prevent the coda filter from applying if the [PLACE] node is linked to more than one consonant. If we assume that syllable

one mora. Such a syllable structure, coupled with these definitions of *onset* and *coda* will raise problems for the approach to autosegmental licensing discussed below. However, the light closed syllable is only rarely discussed (e.g. Hayes 1989) and even less often used in an analysis. I shall leave this matter open for now, as it is of relatively minor relevance to the following discussion.

²⁵ This diagram contains no omissions. It is a faithful reproduction from Itô (1989) example (5).

codas cannot have more than one consonant and that the [PLACE] node can only be linked to *adjacent* C slots, then a doubly-linked [PLACE] node will be linked to a syllable onset. In other words, the coda filter encodes the observation that, regardless of what else a [PLACE] node is linked to, it must be linked to an onset. This amounts to a circuitous version of the expression in (3–43).

Prosodic Licensing and Stray Erasure. This notion, as distinct from autosegmental licensing, “requires that all phonological units belong to higher prosodic structure: segments to syllables, syllables to metrical feet, and metrical feet to phonological words or phrases” (Itô 1989:220). Presumably the prosodic hierarchy is finite, and so there must be some level which is not required to be prosodically licensed. We shall take this to be the phrase level. Now prosodic licensing may be defined as follows.

(3–45) **Prosodic Licensing:**

For all nodes P, either P is of the sort *phrase* or there is a node Q of sort *phrase* where Q dominates P.

This states that all phonological units are either phrases or are dominated by phrases. The sort constraints will ensure that, for example, syllables are dominated by feet and not moras or anything else. It is only necessary to state that everything is ultimately dominated by a phrase. However this formulation is too strong, since it is often required that unattached material be deleted rather than linked (‘stray erasure’). Therefore, rather than adopting (3–45), we only give a phonetic interpretation to connected structures, whereby the denotation of a phonological unit will contain denotations of its constituents only²⁶.

At this point we leave the development of the phonological theory, and turn to its formalization in terms of the theoretical foundation established in chapter two and the language \mathcal{L} , along with its abbreviatory devices.

²⁶ The formalization of this notion remains an open question. While it is quite natural to view unconnected material as having a temporally indeterminate denotation, it is more difficult to come up with a logical basis for this proposal.

3.4 Formalizing the Theory

In this section, I show how the proposals of the last two sections might be formalized using the language \mathcal{L} of chapter two. The presentation here is rather terse, because the motivation and explanation of each definition or constraint has already been given in detail.

The statements in (3–46) define the hierarchical structure. Recall that the parenthesised numbers indicate the permissible degree of branching; the numbers given here are only tentative.

- (3–46) a. $\text{foot} \Rightarrow [\text{syl}(1,3)]$
b. $\text{syl} \Rightarrow [\text{mora}(1,2)]$
c. $\text{mora} \Rightarrow [\text{oral}(0,2), \text{velum}(0,1), \text{larynx}(0,1)]$
d. $\text{oral} \Rightarrow [\text{lip}(0,1), \text{tongue}(0,2)]$
e. $\text{tongue} \Rightarrow [\text{tip}(0,2), \text{body}(0,2)]$

Each of the (active) articulators has a constriction location.

- (3–47) **Constriction Locations:**
a. $\text{lips} = \{\text{protruded}, \text{labial}, \text{labio-dental}\}$
b. $\text{tip} = \{\text{linguo-labial}, \text{dental}, \text{alveolar}, \text{post-alveolar}, \text{palatal}\}$
c. $\text{body} = \{\text{dorso-palatal}, \text{velar}, \text{uvular}, \text{pharyngeal}\}$
d. $\text{velum} = \{\text{velar}\}$
e. $\text{larynx} = \{\text{raised}, \text{lowered}\}$

All five active articulators have a constriction degree. The hierarchical nodes also have constriction degrees.

- (3–48) **Constriction Degrees:**
a. $\text{lips} = \{\text{l-closure}, \text{l-critical}, \text{l-narrow}, \text{l-mid}, \text{l-wide}\}$
b. $\text{tip} = \{\text{t-closure}, \text{t-critical}, \text{t-narrow}, \text{t-mid}, \text{t-wide}\}$
c. $\text{body} = \{\text{b-closure}, \text{b-critical}, \text{b-narrow}, \text{b-mid}, \text{b-wide}\}$
d. $\text{velum} = \{\text{v-closure}, \text{v-wide}\}$
e. $\text{larynx} = \{\text{g-closure}, \text{g-critical}, \text{g-wide}\}$
f. $\text{tongue} = \{\text{closure}, \text{critical}, \text{narrow}, \text{mid}, \text{wide}\}$
g. $\text{oral} = \{\text{o-closure}, \text{o-critical}, \text{o-narrow}, \text{o-mid}, \text{o-wide}\}$
h. $\text{mora} = \{\text{m-closure}, \text{m-critical}, \text{m-narrow}, \text{m-mid}, \text{m-wide}\}$

The classification of constriction degrees into ‘open’ and ‘not open’ which was used in the reconstruction of natural classes in (3–24) is formalized as follows: ‘open’ =

{narrow, mid, wide}, 'not open' = {closure, critical}. This statement pertains only to tongue constriction degrees; similar statements must be made for each of the other articulators.

Furthermore, some nodes can have sequences of constriction degrees. For now I shall assume that tongue and oral nodes can have two constriction degrees, and that mora nodes can have up to three constriction degrees. Multiple constriction degrees are represented as ordered n-tuples. (In (3–49), the constriction degrees are abbreviated as follows: cl(asure), cr(itical), n(arrow), m(id), w(ide).)

- (3–49) a. tongue = { (cl, cl), (cl, cr), (cl, n), (cl, m), (cl, w), (cr, cl), (cr, cr), ... }
 b. oral = { (o-cl, o-cl), (o-cl, o-cr), (o-cl, o-n), (o-cl, o-m), (o-cl, o-w), ... }
 c. mora = { (m-cl, m-cr, m-n), (m-cl, m-cr, m-m), (m-cl, m-n, m-m), ... }

The above properties do not prevent a labio-dental articulation coinciding with a linguo-labial articulation, and so a further constraint is needed. (Note that this is the only such constraint that is necessary, assuming the possible mappings between articulators and constriction locations suggested by Browman & Goldstein 1989:227).

(3–50) **Oral No-Crossing Constraint:**

$$\forall xy \ x \circ y \rightarrow (\text{labio-dental}(x) \rightarrow \neg \text{linguo-labial}(y))$$

Sorts will also be used to represent articulator shape distinctions, e.g., for lateral versus central tongue shape and round versus spread lip shape.

- (3–51) a. tongue = {central, lateral}
 b. tip = {t-central, t-lateral}
 c. body = {b-central, b-lateral}
 d. lips = {round, spread}

Moras and syllables can be further classified.

- (3–52) a. syl = { open, closed }
 b. syl = { light, heavy }
 c. mora = { c-mora, v-mora, cv-mora, cvc-mora }
 d. mora = { onset, coda }

Now the shape of syllables can be given in terms of articulations, according to the following axioms:

(3–53) **Syllable Shape Constraints:**

- a. $\forall x \text{ heavy}(x) \wedge \text{closed}(x) \rightarrow \exists yz \text{ mora}(x, \langle y, z \rangle) \wedge \text{cv-mora}(y) \wedge \text{c-mora}(z)$
- b. $\forall x \text{ heavy}(x) \wedge \text{open}(x) \rightarrow \exists yz \text{ mora}(x, \langle y, z \rangle) \wedge \text{cv-mora}(y) \wedge \text{v-mora}(z)$
- c. $\forall x \text{ light}(x) \wedge \text{closed}(x) \rightarrow \exists y \text{ mora}(x, y) \wedge \text{cvc-mora}(y)$
- d. $\forall x \text{ light}(x) \wedge \text{open}(x) \rightarrow \exists y \text{ mora}(x, y) \wedge (\text{v-mora}(y) \vee \text{cv-mora}(y))$

As the information about headedness is purely relational, headedness will be effectively expressed as a property of the dominance relation. A binary relation $\delta_h (\subseteq \delta)$ and its transitive closure $\delta_h^* (\subseteq \delta^*)$ will be adopted for expressing headedness. I shall adopt a further abbreviatory device, as exemplified in (3–54).

(3–54) **Abbreviatory Device 7:**

- a. $x \delta_h y \wedge \text{tongue}(x) \wedge \text{tip}(y)$
- b. $\text{tip}_h(x, y)$

Now we can employ axioms of the following form for constriction degree percolation.

- (3–55) a. $\forall xyz \text{ tip}_h(x, \langle y, z \rangle) \wedge (\text{cl}, \text{cr})(x) \rightarrow \text{t-cl}(y) \wedge \text{t-cr}(z)$
 b. $\forall xyz \text{ oral}_h(x, \langle y, z \rangle) \wedge (\text{m-cl}, \text{m-n}, \text{m-w})(x) \rightarrow ((\text{o-cl}, \text{o-n})(y) \wedge \text{o-w}(z) \vee \text{o-cl}(y) \wedge (\text{o-n}, \text{o-w})(z))$

Given that constraints of this kind will be stated for each sort of node a given node can dominate, the requirement for multiple inheritance of a single constriction degree to nodes of different sorts is automatically achieved. (Note, however, that this leaves open the question of how to formalize the percolation of *closure* (for English) to only one constituent.)

Autosegmental licensing is expressed as follows:

(3–56) **Autosegmental Licensing:**

$$s_1 \text{ licenses } s_2 =_{\text{def}} \forall x s_2(x) \rightarrow \exists y (s_1(y) \wedge y \delta^* x)$$

This definition may be paraphrased as follows: “a prosodic sort s_1 (e.g. syllable, mora etc.) *licenses* a prosodic sort s_2 if for all nodes of sort s_2 there is a node of sort s_1 dominating s_2 ”.

The connectedness property, if required, can be expressed as in (3–57). Here I am assuming that *phrase* is the highest level category.

- (3–57) $\forall x \exists y \text{ phrase}(x) \vee \text{ phrase}(y) \wedge y \delta^* x$

3.5 Three Kinds of Phonological Generalization

Up to this point we have seen some details about the representation of so-called 'phonological structures', using the constraint language \mathcal{L} . Various kinds of phonological 'rules', or 'alternations', or what I shall call *generalizations*, are also to be cast in this language. Phonologists have traditionally distinguished at least three kinds of such generalizations, and I shall discuss each briefly in turn.

Phonetic structure constraints, or phonotactics. Many generalizations which can be made about sound structure ultimately come down to the organization of the vocal tract, the physics of speech and other aspects of the language faculty with which all humans are endowed. Other generalizations are language or dialect specific (such as the pronunciation of *r* or the presence of intrusive stops (Fourakis & Port 1986, Clements 1987)). Both of these kinds of generalizations concern audible distinctions which, for one reason or another, are not made in a particular language. The present approach relies heavily on the notion of partiality, whereby the precise temporal organization and featural content of a linguistic description may not be fully determined. While some of this partiality may ultimately be manifested as so-called 'free variation', most of the time further constraints (either language particular or universal) are needed to fill in the remaining details. It is likely that the observations about fine temporal coordination will require a richer set of temporal relations than the \circ and \prec relations or reference to a richer array of gestural events. For example, for the production of an aspirated oral stop, a glottalic widening gesture x must overlap *and finish after* an oral closure gesture y . One possibility would be to write x overlaps-and-finishes-after y , and provide axioms for this temporal relation. A second possibility would be to refer to the start and finish of a gesture²⁷ and state that the oral closure finish precedes (i.e. \prec) the glottal widening finish.

'Unnatural' or 'Truly Phonological' Generalizations. Partial specification in phonological theories is ultimately just a special kind of disjunction. This is because the background information about the appropriateness of certain values for certain fea-

²⁷ The start and finish of a gesture do not correspond to the endpoints of the temporal interval of that gesture but rather to intervals (which may be small).

tures is disjunctive. For example, a vowel which is unspecified for a value of a feature such as [high] could equally well have been specified $[+high] \vee [-high]$. Generalizations based on partial specification are often called ‘natural’ due to their close links with phonetic structure. However, many alternations are not of this variety. For these we can still use disjunction, but of a more general variety²⁸. For example, the [t] – [ʃ] alternation in pairs like *permit* – *permission* involves the alternation of an alveolar stop and a palato-alveolar fricative. A possible analysis would involve a lexical representation of the root where the last ‘segment’ (i.e. mora node or oral node) is described disjunctively as in (3–58).

$$(3-58) \text{ mora}(a) \wedge a \delta^* b \wedge \text{tip}(b) \wedge (\text{closure}(a) \wedge \text{alveolar}(b) \vee \text{critical}(a) \wedge \text{palato-alveolar}(b))$$

To this we can conjoin the default specification in (3–59) which gives us [t] as the ‘isolation form’.

$$(3-59) \text{ M } (\text{closure}(a) \wedge \text{alveolar}(b)) \rightarrow \text{closure}(a) \wedge \text{alveolar}(b)$$

Now, suppose that a formula $\phi(x)$ is true just in case x is a mora which contains [i], and suppose $\psi(x,y)$ is true just in case x is a mora which contains an instance of the [t] – [ʃ] alternation involving node y . Then the constraint in (3–60) gives us the [ʃ] variant if the vowel of the mora is [i].

$$(3-60) \forall xy \phi(x) \wedge \psi(x,y) \rightarrow \text{critical}(x) \wedge \text{palato-alveolar}(y)$$

A similar approach can be employed to alternations which involve different articulators (such as the [k] – [s] alternation in *electric* – *electricity*). The morphological sensitivity required for some of these alternations can be achieved, I believe, by making constraints like the one in (3–60) part of the description of the conditioning morpheme itself. However, the details remain the subject of further investigation.

Morpheme structure conditions and morphologically sensitive alternations. Kisseberth (1970) showed how a derivational account of Yawelmani phonology missed important generalizations about the size of consonant clusters in various word positions. Although he proposed a tentative solution using “derivational constraints” to the

²⁸ This approach employs disjunctions of arbitrary sorts. I anticipate that an algebraic approach would begin by generalizing the sort lattices of chapter two to boolean lattices.

problem of rule-relatedness, the issue seems to have been by-and-large passed over in subsequent literature. Nevertheless, it seems to be unavoidable that derivationally expressed generalizations about syllable structure contain rule conspiracies and thereby miss generalizations. My tentative conclusion—which cannot be adequately defended here—is that the widespread assumption that phonological structures in the lexicon completely lack syllable structure is invalid. Instead, I propose that the lexical forms of morphemes have partially specified prosodic structure, and that generalizations about this structure be expressed as *lexical generalizations*. For example, in a language where only syllable onsets license place of articulation specification, each such specification will be linked *lexically* to an onset node. The generalization will then be that each lexical entry containing a place of articulation specification will also have an onset node dominating that specification. I expect such generalizations to be formalized as part of the sort system: linguistic signs, each having phonology, syntax and semantics attributes, will be assigned to various classes depending upon their grammatical status. Signs which are morphologically simple will have the sort *lexical-sign*, the rest have the sort *phrasal-sign* (following Pollard & Sag 1987:39ff). Therefore, a lexical generalization will have the form given in (3–61), where $\phi(x)$ is some phonological constraint (such as the one we saw in section 4 for licensing).

$$(3-61) \quad \forall x \text{ lexical-sign}(x) \wedge \text{phon}(x,y) \rightarrow \phi(y)$$

If $\phi(x)$ were to apply pan-derivationally (rather than just lexically), then we could drop the condition that x be a *lexical-sign* and permit it to be any kind of sign. In this way, the frequently expressed concerns about duplication between morpheme structure constraints and surface phonotactic constraints can be addressed, as it is possible to generalize across both kinds of constraint. Further discussion of this view of grammatical organization goes beyond the scope of the present discussion. An important caveat is that this view is yet to be squared with the frequently advocated position that non-root morphemes cannot be domains of rule application (e.g. Poser 1989:120).

3.6 Conclusion

In this chapter I have attempted to apply the formalism of chapter two to a theory of phonological structure. The literature on feature geometry was surveyed, and the connection between the extant proposals for hierarchical organization of phonological structures and the physiological structure of the vocal tract was observed. This observation has previously been exploited by Sagey, Browman and Goldstein, and an attempt was made to bring both proposals together by identifying Sagey's arrow notation with Browman and Goldstein's constriction degree percolation mechanism. The resulting theory also encompassed some important observations about autosegmental and prosodic licensing. Finally, this theory was formalized using the framework of chapter two. Throughout, I have maintained a clear distinction between the linguistic theory and its formalization, and a clear distinction between the language(s) being analysed and the language \mathcal{L} in which the linguistic generalizations are expressed. The last step is to provide the formalism with a procedural interpretation so that various kinds of computation (such as generation and recognition) can be performed. This is the topic of the next chapter.

Chapter 4

Implementation

The phonological theories which have grown up around (Chomsky & Halle 1968), known collectively as 'generative phonology', have been built on the computational metaphor known as SYMBOL PROCESSING. This model involves the fundamental division between data and process, which phonologists employ as their representations and rules respectively. However, since that time, other computational metaphors have been developed. For example, the view of computation as NEURAL PROCESSING has recently been advocated in the phonology literature (e.g. Lathroum 1989). The implementation discussed here is based on yet another metaphor. How, then, is one to come up with a theory of phonology which makes no commitment to any of these computational issues?

In the computer science literature it has long been maintained that the statement of a problem and the description of its solution should be expressed in a way that is not sensitive to implementation-specific details. For example, a Prolog programmer using a VAX should not have to change her program to make it work on a SUN. Similarly, the definition of the relation: 'x is the square of y' should not depend on whether squares are to be computed using a Turing machine or a neural network. In an analogous way, then, the statement of a phonological theory should ideally not be tied to a particular strategy for computing phonological well-formedness.

The answer to the above question is, I believe, to adopt the distinction between DENOTATIONAL and OPERATIONAL semantics (as widely practiced in the literature on the principles of programming languages).

A program can be interpreted as an abstract specification of the solution to a certain problem. Under this interpretation, a program denotes a space of solutions. A specification of the interpretation of a program solely in terms of the solution space it describes is called a denotational semantics. In contrast, the operational semantics of a program specifies how it can be interpreted as an algorithm for computing the set of solutions. There can be more than one operational semantics for a program, depending upon the kind of computing machinery to be used for the computation. For example, the interpretation of a program as a set of instructions for a Turing machine is rather different to the interpretation of the same program for a neural network, even though the solutions which are computed will be the same. Therefore, in order to avoid any bias towards a particular computational metaphor, it is best that programming languages be first endowed with a denotational semantics, and a variety of operational semantics can follow later. This distinction in interpretation is sometimes called the declarative/procedural distinction¹.

Theories of linguistics almost invariably provide the practitioner with a descriptive vocabulary and a means of formulating observations about natural languages using that vocabulary. This, then, is what might be called a 'language for linguistic description'. As already discussed in chapter one section one, such a language has a syntax and a semantics. A syntax specifies which expressions are well-formed expressions of the language. A semantics specifies how these expressions correspond to objects in the domain being described. In chapter two section eight the syntax of a phonological description language \mathcal{L} was provided. The language is a language of first order logic, and so the existence of a denotational semantics is guaranteed. There are a variety of ways in which \mathcal{L} might be given an operational semantics. In

¹ I am only aware of one instance in the phonology literature where the declarative/procedural distinction is mentioned. In discussing "declarative rules" versus "procedural rules", Bromberger & Halle (1989) state in a footnote: "[This] terminology, which carries a number of associations from the domain of computational linguistics, strikes us as unhelpful." Unfortunately they give no justification for this statement. Here I demonstrate some of the benefits of incorporating the declarative/procedural distinction and its associations with the domain of computational linguistics into phonological thinking.

what follows, I shall give an informal presentation of just one of them.

The semantics I shall describe is based around a computational metaphor known as DEDUCTIVE INFERENCE². Each computation step involves an application of an inference rule (such as modus ponens). With each step, a new piece of information is inferred. As long as each new piece of information is consistent with the existing information the inferencing process can continue until no more rules of inference can be applied. If a new piece of information contradicts the existing information then the inferencing process cannot continue any further. This computational technique can be used to test the consistency of an expression of \mathcal{L} : we simply compute all of the logical consequences of the expression and check that no contradictions can be derived.

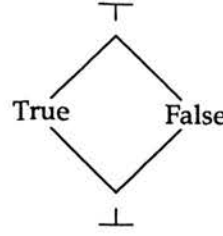
In this chapter I give an informal description of an operational semantics for \mathcal{L} . First, the representation of atomic formulas and axioms is given. Next, the inferencing process which manipulates these representations is described. Finally, a small example is provided to illustrate the functioning of the system. A pilot version of the system has been implemented using the Prolog and C languages.

4.1 Representation

Information about the sorts, featural and temporal relations is maintained in arrays, indexed by the (finite) set of constant symbols. This follows the strategy used by Allen (1983) and Schmiedel (1988). For each sort predicate s there is a one-dimensional array A_s , and for each featural and temporal relation r there is a two-dimensional array A_r . For each constant symbol c there is an index i_c . Each array cell encodes a truth value from a four-valued logic:

² This view has most recently been exploited within the paradigm of CONSTRAINT LOGIC PROGRAMMING (Jaffar & Lassez 1987, Smolka *to appear*).

(4-1)



If $r(a, b)$ is true (false) then $A_r[i_a, i_b]$ is set to True (False). If $r(a, b)$ is unknown then the corresponding cell is set to an expression meaning 'True or False', and if its value is inconsistent the cell is set to an expression meaning 'True and False'. Internally, these truth values are stored using bit strings. For each cell there are two bits, the first encoding 'possible truth' and the second encoding 'possible falsehood'. The correspondence between the truth values and the bit strings is according to the following table:

(4-2)

Truth Value	Unknown	True	False	Inconsistent
Bit String	11	10	01	00

If we begin with the information that $r(a, b)$ is true (i.e. 10) and later add the information that $r(a, b)$ is false (i.e. 01), we do not simply replace 10 with 01. Instead we must perform the bitwise-and operation between the two values, giving the result 00. This is the desired result, because an expression which implies that $r(a, b)$ is simultaneously true and false is inconsistent.

Axioms will typically have the form $r(x, y) \rightarrow r'(x, y)$. These are therefore general constraints between the entries of different tables. Each time a table entry is updated, these axioms can be used to infer further information and update other table entries, and so on. For technical reasons, the system is restricted to the class of universally-quantified function-free prenex formulae with equality³. Each axiom is translated into a formula in conjunctive normal form:

$$(4-3) \quad \forall x_1 \dots x_n (l_1 \vee \dots \vee l_m) \wedge \dots \wedge (l'_1 \vee \dots \vee l'_m)$$

In (4-3), each l is a (possibly negated) predicate (i.e. a literal). The next step is to convert each conjunct into a new axiom. Consequently, all axioms have the form in (4-4).

³ Note that the restriction on quantification can be removed trivially, given the fact that there are only a finite number of constants.

$$(4-4) \quad \forall x_1 \dots x_n \quad l_1 \vee \dots \vee l_m$$

The final step translates each axiom into a piece of C code, the details of which will not be of any concern here. Instead, some of the axioms from chapter two and chapter three are presented in the above format.

Background Axioms

For all x, y and z :

1. $x \circ x$
2. $\neg x \circ y \vee y \circ x$
3. $\neg x \prec y \vee \neg y \prec x$
4. $\neg x \prec y \vee \neg x \circ y$
5. $\neg x \prec y \vee \neg y \prec z \vee x \prec z$
6. $x \prec y \vee x \circ y \vee x \prec y$
7. $\neg x \prec y \vee \neg y \circ z \vee \neg z \prec x$
8. $\neg x \delta x$
9. $\neg x \delta y \vee \neg y \delta x$
10. $\neg x \delta^* y \vee x \circ y$

General Equality Axioms:

1. $x = x$
2. $\neg x = y \vee y = x$
3. $\neg x = y \vee \neg y = z \vee x = z$
4. $\neg x = y \vee \neg y \circ z \vee x \circ z$
5. $\neg x \circ y \vee \neg y = z \vee x \circ z$
6. $\neg x = y \vee \neg y \prec z \vee x \prec z$
7. $\neg x \prec y \vee \neg y = z \vee x \prec z$

$$8. \neg x = y \vee \neg y \delta z \vee x \delta z$$

$$9. \neg x \delta y \vee \neg y = z \vee x \delta z$$

Linguistic Axioms

The following axioms are based on chapter 3 section 4. This is not a complete set but an illustrative selection. The complete set of axioms from chapter 3 in unabbreviated form would run for pages.

$$1. \neg \text{foot}(x) \vee \neg x \delta y \vee \text{syl}(y)$$

$$2. \neg \text{syl}(x) \vee \neg x \delta y \vee \text{mora}(y)$$

$$3. \neg \text{mora}(x) \vee \neg x \delta y \vee \text{oral}(y) \vee \text{velum}(y) \vee \text{larynx}(y)$$

$$4. \neg \text{oral}(x) \vee \neg x \delta y \vee \text{lip}(y) \vee \text{tongue}(y)$$

$$5. \neg \text{tongue}(x) \vee \neg x \delta y \vee \text{tip}(y) \vee \text{body}(y)$$

$$6. \neg \text{lip}(x) \vee \text{protruded}(x) \vee \text{labial}(x) \vee \text{labio-dental}(x)$$

$$7. \neg \text{labial}(x) \vee \text{lip}(x)$$

$$8. \neg \text{lip}(x) \vee \text{l-closure}(x) \vee \text{l-critical}(x) \vee \text{l-narrow}(x) \vee \text{l-mid}(x) \vee \text{l-wide}(x)$$

$$9. \neg \text{lip}(x) \vee \text{round}(x) \vee \text{spread}(x)$$

4.2 An Example

In this section a small example is given to illustrate the workings of the system. (This particular example is designed to show how the no-crossing constraint is implemented.)

Suppose we know that there is some w, x, y and z such that $w \prec x, x \circ y$ and $y \prec z$. The tables for \prec and \circ are as follows:

(4-5) a.	A_{\prec}	w	x	y	z
	w	11	10	11	11
	x	11	11	11	11
	y	11	11	11	10
	z	11	11	11	11
b.	A_{\circ}	w	x	y	z
	w	11	11	11	11
	x	11	11	10	11
	y	11	11	11	11
	z	11	11	11	11

Using the above axioms, we are able to infer some positive information:

- (4-6) a. $w \circ w$ (from Background Axiom 1)
b. $x \circ x$ (from 1)
c. $y \circ y$ (from 1)
d. $z \circ z$ (from 1)
e. $y \circ x$ (from 2)
f. $w \prec z$ (from 6, 7)

and a large amount of negative information:

- (4-7) a. $\neg w \prec w$ (from 3)
b. $\neg x \prec x$ (from 3)
c. $\neg y \prec y$ (from 3)
d. $\neg z \prec z$ (from 3)
e. $\neg x \prec w$ (from 3)
f. $\neg x \circ w$ (from 4)
g. $\neg w \circ x$ (from 4, 2)
h. $\neg x \prec y$ (from 4)
i. $\neg y \prec x$ (from 4, 3)
j. $\neg z \prec y$ (from 3)
k. $\neg z \circ y$ (from 4)
l. $\neg y \circ z$ (from 4, 2)
m. $\neg y \prec w$ (from 6)
n. $\neg z \prec x$ (from 6)
o. $\neg z \prec w$ (from 6, 7, 3)
p. $\neg z \circ w$ (from 6, 7, 4)
q. $\neg w \circ z$ (from 6, 7, 4, 2)

This results in the tables being updated as follows:

(4-8) a.	A_{\prec}	w	x	y	z
	w	01	10	11	10
	x	01	01	01	11
	y	01	01	01	10
	z	01	01	01	01
b.	A_{\circ}	w	x	y	z
	w	10	01	11	01
	x	01	10	10	11
	y	11	10	10	01
	z	01	11	01	10

Notice that very few table entries now contain 11. The cases where 11 remains express the following information: w may precede or overlap y , and x may precede or overlap z . Suppose that we now wished to add the information that $w \circ z$. The value in the fourth column of the first row of (4-8b) (i.e. 01) must be updated with the new value (10). These two values must be ‘anded’ together, with the result being 00. This cell now expresses the information that $w \circ z$ and $\neg w \circ z$, which is inconsistent. Consequently, the attempt to add the information $w \circ z$ fails. In phonological terms, this means that the attempt to create a crossing line situation is blocked.

We have now seen an example of the representation and manipulation of *temporal* information and constraints. The representation and manipulation of *linguistic* information and constraints is entirely analogous. A table is used to represent the δ relation, and one-dimensional tables represent the sorts (e.g. syl, mora, labial). These tables are updated with information about phonological structures, and the axioms are employed to ensure their consistency.

4.3 Conclusion

A word is in order about the computational complexity of the inferencing process. In the worst case, the complexity may be exponential because each axiom can produce an update which triggers the application of several further axioms. However, due to the relatively small size of most phonological representations (i.e. morpheme size), this complexity may not be a problem. Nevertheless, it could be argued that the reason phonological representations are typically rather small is due to the difficulty

of depicting large multidimensional objects on the page. Now that a computational representation exists, such restrictions no longer need apply. The phonological representation of a whole discourse could be built out of representations of each morpheme, potentially involving many thousands of nodes. In this situation one could reasonably expect a combinatorial explosion to prevent any useful computation from taking place within a reasonable time period. A way around this problem has already been suggested by Allen (1983) and it involves the use of what he calls REFERENCE INTERVALS. "Reference intervals are used to group together clusters of intervals for which the temporal constraints between each pair of intervals in the cluster is fully computed. Such a cluster is related to the rest of the intervals in the system only indirectly via the reference interval" (Allen 1983:838ff). So, for example, if all phonological phrases were reference intervals⁴, then it would not be necessary to explicitly infer that all the constituents of some phrase temporally precede all the constituents of all later phrases. More generally, the use of reference intervals means that the consequence of doubling the size of a phonological structure doubles (rather than squares) the amount of computation required. At this early stage then, there seem to be good reasons to hope that it is possible to sufficiently constrain the proliferation of temporal inferencing. The guarantee that the inferencing process terminates follows from the fact that each update to the tables always reduces the total number of *is* by one.

In its present form, the system is interfaced to a Prolog interpreter. A distinguished Prolog predicate is linked to the main C routine. When the predicate is tested by Prolog, the corresponding C routine is invoked to update a table entry and perform inferences⁵. Once the C routine has finished its work, the success or failure of the update is communicated back to the Prolog interpreter. There are two important consequences of this design. First, the presence of inconsistency is detected at the earliest possible point. Therefore the search for solutions being conducted by the

⁴ For a direct application of Allen's proposal here, we would need to assume that no phonological constituents could be ambiphrasal. In other words, there must be a unique phrase dominating each sub-phrasal constituent. This assumption seems reasonable given the rarity (or perhaps even the non-existence) of sandhi across phrase boundaries.

⁵ The interface between Prolog and C is implemented in a way that respects Prolog's use of choice points and backtracking, making it possible to write pure (i.e. logical) programs. However, the details of this aspect of the implementation are beyond the scope of the present discussion.

Prolog interpreter does not enter into any 'blind alleys', and this has important consequences for efficiency. The second important consequence of the design is that the programmer does not need to be aware of any of the implementation details discussed above. To the programmer it appears as if the Prolog interpreter has a built-in phonological theorem prover.

In conclusion, this chapter has given an informal outline of an operational semantics for the phonological description language presented in chapter two, based on the view of computation as deductive inference. The implemented system has been used to test some of the theoretical proposals advanced earlier.

Chapter 5

Conclusion

The primary goal of this study has been to provide a formal foundation for phonological description that is compatible with existing constraint-based approaches to natural language syntax and semantics. The main theme of the study has been the distinguishing of structures from descriptions.

The formal basis of this work lies in the marriage of temporal logic and feature logic. This novel move was motivated and developed in chapter 2, on the basis of a range of proposals (both suprasegmental and subsegmental) from the phonology literature. The formalism was couched in classical first order logic because work by van Benthem (1983) on temporal logic and Johnson (1990) on feature logic has shown the fruitfulness of such an approach. Notably, the interaction between temporal and featural predicates was easy to explore in a first order framework. This formally motivated proposal for adopting feature structures in phonology is complemented by Hayes' (1990) phonologically motivated proposal for the adoption of feature structures (reviewed in Appendix B). The formalism is offered as a first step towards a uniform logical foundation for all levels of linguistic description, including the syntactic, the semantic and the phonological.

With this as a basis, a model of phonological structure was proposed in chapter 3. Two important recent developments in phonology, one concerning abstract phonol-

ogy (Sagey 1986) and the other concerning the phonology-phonetics interface (Browman & Goldstein 1989), were brought together. Sagey used phonological argumentation to defend a view of hierarchical feature organization grounded in articulatory phonetics. She also proposed an arrow notation which expressed the relationship between manner features and place features. I argued that this notation can be viewed as expressing a relationship between acoustic and articulatory properties. This view was shown to be closely related to Browman & Goldstein's (1989) use of 'constriction degree percolation', whereby the net constriction degree of the vocal tract is related to the constriction degrees of the various parts of the vocal tract. The resulting model of hierarchical feature organization was formalized in terms of the framework developed in chapter 2. Finally, chapter 4 provided a brief overview of a computational implementation in the paradigm of constraint logic programming.

This development has a number of consequences for theoretical phonology. First, phonological representations can be described succinctly, without recourse to diagrams or prose, although I do not deny that these latter have their place; indeed the opportunities for visualization offered by an apt graphical notation are invaluable. Second, inference and consistency checking can be performed directly on representations. Third, abbreviatory devices can be stated formally, with the consequence that using abbreviated forms is less likely to lead to confusion. Fourth, theoretical claims can be made more explicit if their substantive content is unambiguous, allowing a formal comparison of competing analyses to be made. Finally, the resulting model—being declarative—provides for a non-derivational view of phonology (following in the footsteps of Hooper 1976 and Hudson 1980, 1986). This, in turn, enables generalizations about linguistic competence to be stated independently of particular performance tasks, such as generation (the subject of 'generative phonology') or recognition (the subject of 'upside-down phonology', Leben & Robinson 1977).

There are also consequences for descriptive phonology. The formalism provides a computational representation for phonological descriptions. This development will ultimately enable the automatic checking of the correspondence between an analy-

sis and its target data. Gestural scores—the lowest level of representation employed here—have been used successfully as the basis for automatic speech synthesis (Browman pers. comm.). In elaborating the relationship between articulatory phonetic structure and a particular model of abstract phonology, a practical (yet theoretically well-founded) view of the phonology-phonetics interface is emerging, one that will ultimately enable the testing of phonological grammars using speech synthesis.

A third set of consequences arise for computational linguistics. Although a variety of constraint-based grammar formalisms have been presented in the literature, they have hitherto been employed mainly for analyses of syntactic and semantic phenomena. Applications to morphological and phonological domains have been severely curtailed because these formalisms have assumed an overly restrictive view of phonological organization, whereby representations are conceived of as strings over an alphabet. While certain string manipulations can account for a variety of phenomena (e.g. Hoeksema & Janda 1988), the wholesale conflation of phonology with orthography renders any theory incapable of expressing the observations which have been made in the non-linear phonology literature. Once these formalisms have been suitably enriched, perhaps along the lines of the present proposals, it will be possible to study a wider class of languages, particularly those with complex morphophonologies (such as the Bantu and Semitic families, to name but two).

Discussion of classes of languages immediately raises the question about the weak generative capacity of the grammars which can be encoded in \mathcal{L} . However, since \mathcal{L} is not a system of production rules which licences strings of alphabetic symbols, such questions are ill-posed. Furthermore, since there is ultimately more to phonological structures than their terminal yield¹, the standard distinction between weak and strong generative capacity loses much of its force. The source of the difficulty in applying generative capacity measurements is simply this: the languages we wish to describe are not linear but are multi-linear. For each articulator of the vocal tract there is a linear sequence of specifications. With multiple articulators and a complex

¹For example, in the moraic theory of syllable structure, the distinction between light and heavy syllables is expressed using moras (which correspond to non-terminal symbols). It is possible to have two syllables with the same terminal yield where one syllable is heavy and the other is light.

set of constraints on coarticulation, there is no obvious way in which the standard definitions from formal language theory can be carried over.

A final comment about the association relation is in order. This relation, which lies at the heart of non-linear approaches to phonology, has been formalized as a temporal overlap relation, following Sagey. However, a novel view of its function has emerged from the prosodic view taken here. As everything is ultimately a part of prosodic structure, the generalizations about feature occurrence are to be stated in terms of prosodic categories.

For example, on the standard view, if all the vowels of a word are to agree in height, there is a separate tier for the feature [high] and a rule which spreads a single [+high] or [-high] specification to all V-slots. The present view is radically different. Here, a high vowel is just a *partially specified syllable* containing the mora: [deg narrow]_{mora}. A harmony rule would be expressed as the following constraint: the syllables of a word must all satisfy ϕ , or all must satisfy ψ , or all must satisfy χ , where ϕ , ψ and χ are partial descriptions of syllables (corresponding to high, mid and low vowels respectively). A similar view applies to any other property which is traditionally held to 'associate at the syllable level'. Consonant spreading is most likely to be analysed as generalizations over sequences of moras or lower nodes, rather than syllables. The important feature of the formalism is that through the use of composite relations (i.e. the $f|g(x,y)$ notation), we can refer to the sequence of (say) oral nodes of a syllable without any reference to the intervening mora nodes, and no constraint on their number or makeup. Tiers are simply viewed as partial descriptions of prosodic structure. These partial descriptions are combined using conjunction; the association relation does not figure here at all. The latter is reserved as a constraint on the temporal organization of the sister constituents of a node in hierarchical structure.

This proposal for a constraint-based approach to phonology is certain to have raised a good many more questions than it has answered. Indeed, the relative formality of the approach is no guarantee of linguistic or mathematical correctness; on the contrary, its success is to be measured in proportion to the ease with which the theory

can be formally refuted and replaced by a theory with fewer shortcomings. Such a programme will, I hope, ultimately provide us with a theory of phonology that is sufficiently rigorous and precise to be computationally interpretable and testable, an approach to speech technology that is linguistically well-founded and well-integrated with syntax and semantics, and a way of applying extant grammar frameworks such as HPSG to a significantly broader range of the world's languages than has previously been possible.

References

- Allen, James F. (1983). Maintaining knowledge about temporal intervals. *Communications of the ACM* 26, 832–843.
- Anderson, John. (1988). More on slips and syllable structure. *Phonology* 5, 157–159.
- Anderson, Stephen. (1981). Why phonology isn't 'natural'. *Linguistic Inquiry* 12, 493–539.
- Anderson, Stephen. (1985). *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Chicago: The University of Chicago Press.
- Archangeli, Diana. & Douglas Pulleyblank. (1987). Maximal and minimal rules: effects of tier scansion. *NELS* 17, 16–35.
- Árnason, Kristján. (1986). The segmental and suprasegmental status of preaspiration in Modern Icelandic. *Nordic Journal of Linguistics* 9, 1–23.
- Avery, Peter & Keren Rice. (1989). Segment structure and coronal underspecification. *Phonology* 6, 179–200.
- Bach, Emmon. (1976). An extension of classical transformational grammar. *Problems in Linguistic Metatheory, Proceedings of the 1976 Conference at Michigan State University*, 183–224.
- Batóg, Tadeusz. (1967). *The Axiomatic Method in Phonology*. London: Routledge & Kegan Paul.
- Bell-Berti, F. & Katherine S. Harris. (1979). Anticipatory coarticulation: some implications from a study of lip rounding. *Journal of the Acoustical Society of America*

- Bennett, M. & Barbara Partee. (1972). *Toward the Logic of Tense and Aspect in English*. Indiana University Linguistics Club.
- van Benthem, Johan. (1983). *The Logic of Time: A Model-Theoretic Investigation into the Varieties of Temporal Ontology and Temporal Discourse*, Synthese Library 156, Dordrecht: Reidel.
- van Benthem, Johan. (1988). *A Manual of Intensional Logic*, CSLI Lecture Notes, Number 1. Second Edition.
- Bird, Steven. (1990). *Feature Structures and Indices*. ms.
- Bird, Steven & Ewan Klein. (1990). Phonological Events. *Journal of Linguistics* 26, 33–56.
- Bird, Steven & D. Robert Ladd. (forthcoming). Review of Goldsmith 1990. *Journal of Linguistics* 27.
- Bloch, Bernard. (1948). A set of postulates for phonemic analysis. *Language*. 24, 3–46.
- Bloomfield, Leonard. (1926). A set of postulates for the science of language. *Language*. 2, 153–164. Reprinted in Martin Joos (ed.) *Readings in Linguistics I: The Development of Descriptive Linguistics in America 1925–56*. 26–31.
- Broe, Michael. (1988). A unification-based approach to prosodic analysis. In Docherty, Miller & Stirling (eds.) *Work in Progress* 21, 63–82. Dept. of Linguistics, University of Edinburgh.
- Broe, Michael. (to appear). An introduction to feature geometry. In Docherty & Ladd (eds.) *Gesture, Segment, Prosody: Papers in Laboratory Phonology II*. Cambridge University Press.
- Bromberger, Sylvain. (1988). *Types and Tokens in Linguistics*. CLSI-88-125, Stanford University, April 1988.
- Browman, Catherine & Louis Goldstein. (1986). Towards an articulatory phonology. In Colin Ewen & John Anderson (eds.) *Phonology Yearbook*. palb 3, 219–252.
- Browman, Catherine P. & Louis Goldstein. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica* 45, 140–155.

- Browman, Catherine P. & Louis Goldstein. (1989). Articulatory gestures as phonological units. *Phonology* 6, 201–251.
- Calder, Jonathan. (1990). *An Interpretation of Paradigmatic Morphology*. Ph.D. Thesis, University of Edinburgh.
- Calder, Jonathan, Ewan Klein & Henk Zeevat. (1988). Unification Categorical Grammar: A Concise, Extendable Grammar for Natural Language Processing. *COLING 88*, Budapest.
- Catford, John C. (1988). *A Practical Introduction to Phonetics*. Oxford: Clarendon Press.
- Chen, Matthew. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129–159.
- Chomsky, Noam. (1959). On certain formal properties of grammars. *Information and Control* 1, 137–167.
- Chomsky, Noam. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam & Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Clements, George N. (1985). The geometry of phonological features. In Ewen & Anderson (eds.) *Phonology Yearbook* 2, 223–52. Cambridge University Press.
- Clements, George N. (1987). Phonological feature representation and the description of intrusive stops. In Bosch, Need & Schiller (eds.) *23rd Annual Regional Meeting of the Chicago Linguistics Society. Part Two: Parasession on Autosegmental and Metrical Phonology*. Chicago Linguistics Society. 29–50.
- Clements, George N. & Samuel J. Keyser. (1983). *CV Phonology*. Linguistic Inquiry Monograph Series, Number 9. Cambridge, MA: MIT Press.
- Coleman, John S. (1990). Charm theory defines strange vowel sets. *Journal of Linguistics* 26, 165–174.
- Cooper, Richard P. (forthcoming). *Classification-based Phrase Structure Grammar: an Extended Revised Version of HPSG*. Ph.D. Thesis. Centre for Cognitive Science, University of Edinburgh.

- Dalrymple, Mary, Ronald M. Kaplan, Lauri Karttunen, Kimmo Koskenniemi, Sami Shaio & Michael Wescoat. (1987). *Tools for Morphological Analysis*. CSLI-87-108, Stanford University, September 1987.
- Davis, Stuart. (1989). [Continuant] in feature geometry. *Lingua* 78, 1–22.
- Dennett, Daniel C. (1987). *The Intensional Stance*. Cambridge, MA: MIT Press.
- Dinnsen, Daniel A. (1985). A re-examination of phonological neutralization. *Journal of Linguistics* 21, 265–279.
- Dixon, Robert M. W. (1963). *Linguistic Science and Logic*. The Hague: Mouton.
- Dresher, B. Elan. (1981). Abstractness and explanation in phonology. In Hornstein & Lightfoot (eds.) *Explanation in Linguistics*. London: Longman.
- Evans, Roger & Gerald Gazdar. (eds.) (1990). *The DATR Papers: February 1990*. University of Sussex, Cognitive Science Research Reports.
- Ewen, Colin. (1986). Segmental and suprasegmental structure. In Durand, J. (ed.) *Dependency and non-linear phonology*. London: Croom Helm. 203–222.
- Firth, John R. (1948/1957). Sounds and Prosodies. *Transactions of the Philological Society*. Also in *Papers in Linguistics 1934–1951*, London: Oxford University Press, (1957), 121–38.
- Fourakis, Marios & Robert Port. (1986). Stop epenthesis in English. *Journal of Phonetics* 14, 197–221.
- Fowler, Carol. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics* 8, 113–133.
- Fudge, Erik. (1969). Syllables. *Journal of Linguistics* 5, 253–286.
- Fudge, Erik. (1987). Branching structure within the syllable. *Journal of Linguistics* 23, 359–377.
- Gazdar, Gerald. (1981). Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12, 155–184.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag. (1985). *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.

- Gibbon, Dafydd. (1990). Prosodic association by template inheritance. *International Workshop on Inheritance and Natural Language Processing*, Tilburg, The Netherlands, August 1990.
- Goldsmith, John A. (1976). *Autosegmental Phonology*. Doctoral Thesis, Massachusetts Institute of Technology. Garland Publishing, Inc. (1979).
- Goldsmith, John A. (1990). *Autosegmental & Metrical Phonology*. Oxford: Basil Blackwell.
- Greenberg, Joseph H. (1959). An axiomatization of the phonologic aspect of language. In L. Gross (ed.) *Symposium on Sociological Theory*. 437–480. New York: Evanston.
- Griffin, Toby D. (1976). Toward a nonsegmental phonology. *Lingua* 40, 1–20.
- Griffin, Toby D. (1985). *Aspects of Dynamic Phonology*. Amsterdam Studies in the Theory and History of Linguistic Science, Series IV. Current Issues in Linguistic Theory, Volume 37. Amsterdam & Philadelphia: John Benjamins.
- Gunji, Takao. (1986). *Japanese Phrase Structure Grammar*. Dordrecht: Reidel.
- Gussmann, Edmund. (1980). *Studies in Abstract Phonology*. Linguistic Inquiry Monograph, Number 4. Cambridge MA: MIT Press.
- Hall, Tracy Alan. (1989). Lexical Phonology and the distribution of German [ç] and [x]. *Phonology* 6, 1–17.
- Halle, Morris. (1983). On distinctive features and their articulatory implementation. *Natural Language and Linguistic Theory* 1, 91–105.
- Halle, Morris & Jean-Roger Vergnaud. (1980). Three-dimensional phonology. *Journal of Linguistic Research* 1, 83–105.
- Hammond, Michael. (1988). On deriving the well-formedness condition. *Linguistic Inquiry* 19, 319–325.
- Hayes, Bruce. (1986a). Assimilation as spreading in Toba Batak *Linguistic Inquiry* 17, 467–499.
- Hayes, Bruce. (1986b). Inalterability in CV phonology. *Language* 62, 321–351.

- Hayes, Bruce. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20, 253–306.
- Hayes, Bruce. (1990). Diphthongization and coindexing. *Phonology* 7, 31–71.
- Hoberman, Robert D. (1988). Local and Long-Distance Spreading in Semitic Morphology. *Natural Language and Linguistic Theory* 6, 541–549.
- Hockett, Charles. (1955). *A Manual of Phonology*. Baltimore: Waverly Press.
- Hoeksema, Jack. (1985). *Categorial Morphology*. New York: Garland.
- Hoeksema, Jack and Richard Janda. (1988). Implications of process morphology for categorial grammar. In Oehrle, Bach & Wheeler.
- Hooper, Joan. (1976). *An Introduction to Natural Generative Phonology*. Academic Press.
- Hudson, Grover. (1980). Automatic alternations in non-transformational phonology. *Language* 56, 94–125.
- Hudson, Grover. (1986). Arabic root and pattern morphology without tiers. *Journal of Linguistics* 22, 85–122.
- Hyman, Larry M. (1984). On the weightlessness of syllable onsets. In Brugman & Macaulay (eds.) *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*. University of California, Berkeley.
- Itô, Junko. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory*. 7, 217–259.
- Iverson, Gregory K. (1989). On the category Supralaryngeal. *Phonology* 6, 285–303.
- Jaffar, Joxan, & Jean-Louis Lassez. (1987). Constraint Logic Programming. *ACM Symposium on Principles of Programming Languages*. Munich, January 1987.
- Johnson, C. Douglas. (1972). *Formal Aspects of Phonological Description*. The Hague: Mouton.
- Johnson, Mark. (1988). *Attribute-Value Logic and the Theory of Grammar*. CSLI Lecture Notes Number 16, Stanford University.

- Johnson, Mark. (1990). Expressing disjunctive and negative feature constraints with classical first-order logic. *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*. Pittsburgh, June 1990. 173–179.
- Kahn, Daniel. (1976). *Syllable-Based Generalizations in English Phonology*. Ph.D. dissertation, University of Massachusetts. Reproduced by the Indiana University Linguistics Club.
- Kaplan, Ronald M. & Joan Bresnan. (1982). Lexical-Functional Grammar: A formal system for grammatical representation. In Bresnan (ed.), *The Mental Representation of Grammatical Relations*, The MIT Press.
- Karttunen, Lauri. (1986). *Radical Lexicalism*. Report No. CSLI-86-68, December 1986.
- Karttunen, Lauri, Kimmo Koskeniemi & Ronald M. Kaplan. (1987). A compiler for two-level phonological rules. In Dalrymple et al.
- Kasper, Robert T. & Rounds, William C. (1986). A logical semantics for feature structures. *24th Annual Meeting of the Association for Computational Linguistics*, 257–266.
- Kasper, Robert T. & Rounds, William C. (1990). The logic of unification in grammar. *Linguistics and Philosophy* 13, 35–58.
- Keating, Patricia. (1984). Phonetic and phonological representation of stop consonant voicing. *Language* 60, 286–319.
- Kiparsky, Paul. (1968). Linguistic universals and linguistic change. In Bach & Harms (eds.) *Universals in Linguistic Theory*. New York: Holt.
- Kiparsky, Paul. (1973). Elsewhere in phonology. In Anderson & Kiparsky (eds.) *Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston.
- Kisseberth, Charles W. (1970). On the functional unity of phonological rules. *Linguistic Inquiry* 1, 291–306.
- Kisseberth, Charles W. (1984). Digo tonology. In Clements & Goldsmith (eds.) *Autosegmental Studies in Bantu Tone*, 105–182. Dordrecht: Foris.
- Kornai, András. (1989). *A note on phonological events*. ms. Stanford University.

- Ladefoged, Peter. (1971). *Preliminaries to Linguistic Phonetics*. Chicago: Chicago University Press.
- Ladefoged, Peter, Kay Williamson, Ben Elugbe & Sister Ann Angela Uwalaka. (1976). The stops of Owerri Igbo. *Studies in African Linguistics*, Supplement 6, 147–163.
- Larkin, Jill H. & Herbert A. Simon. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11. 65–99.
- Lass, Roger. (1976). *English Phonology and Phonological Theory, Synchronic and Diachronic Studies*. Cambridge University Press.
- Lass, Roger. (1984). *Phonology, An Introduction to Basic Concepts*. Cambridge University Press.
- Leben, William R. (1973). *Suprasegmental Phonology*. Ph.D. dissertation. Massachusetts Institute of Technology.
- Leben, William R. & Orrin W. Robinson. (1977). 'Upside-down' phonology. *Language* 53, 1–20.
- Lindblom, Björn E. F. & Johan E. F. Sundberg. (1971). Acoustical consequences of lip, tongue, jaw and larynx movement. *The Journal of the Acoustical Society of America* 50, 1167–1179.
- Lisker, Leigh. (1986). "Voicing" in English: a catalogue of acoustic features signaling /b/ versus /p/ in trochees. *Language and Speech* 29, 3–11.
- Maier, David & David S. Warren. (1988) *Computing with logic - logic programming with Prolog* Menlo Park, CA: Benjamin/Cummings Publishing Company
- McCarthy, John. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373–413.
- McCarthy, John. (1984). Theoretical consequences of Montañes Vowel Harmony. *Linguistic Inquiry* 15, 291–318.
- McCarthy, John. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry*. 17, 207–263.
- McCarthy, John. (1988). Feature geometry and dependency: a review. *Phonetica*. 45, 84–108.

- McCarthy, John. (1989). Linear order in phonological representation. *Linguistic Inquiry*. 20, 71–99.
- McCarthy, John & Alan Prince. (1989). Prosodic morphology and templatic morphology. In Eid & McCarthy (eds.) *Proceedings of the Second Annual Symposium on Arabic Linguistics*. Amsterdam: John Benjamins.
- Mermelstein, P. (1973). Articulatory model for the study of speech production. *The Journal of the Acoustical Society of America* 53, 1070–1082.
- Mermelstein, P., S Maeda & O. Fujimura. (1971). Description of tongue lip movement in a jaw-based coordinate system. *The Journal of the Acoustical Society of America* 49, 104.
- Minker, J. (ed.) (1987). *Foundations of Deductive Databases and Logic Programming*. Kaufmann.
- Nolan, Francis. (1985). Idiosyncrasy in coarticulation strategies. In Nolan (ed.) *Cambridge Papers in Phonetics and Experimental Linguistics* 4. University of Cambridge, Department of Linguistics.
- Nolan, Francis. (1986). The implications of partial assimilation and incomplete neutralisation. In Hawkins, Sarah (ed.) *Cambridge Papers in Phonetics and Experimental Linguistics* 5. University of Cambridge, Department of Linguistics.
- Oehrle, Richard T. (1988). Multi-dimensional compositional functions as a basis for grammatical analysis. In Oehrle, Bach & Wheeler.
- Oehrle, Richard T., Emmon Bach & Deirdre W. Wheeler. (eds.) (1988). *Categorical Grammars and Natural Language Structures*. Dordrecht: Reidel.
- Paradis, Carole & Jean-François Prunet. (1989). On coronal transparency. *Phonology* 6, 317–348.
- Partee, Barbara H. (1979). Montague grammar and the well-formedness constraint. In Heny & Schnelle (eds.) *Syntax and Semantics 10: Selections from the Third Groningen Round Table*. Academic Press.
- Pereira, Fernando C. N. & David H. D. Warren. (1980). Definite Clause Grammars for Language Analysis — A Survey of the Formalism and a Comparison with

- Augmented Transition Networks. *Artificial Intelligence* 13, 231–278.
- Pierrehumbert, Janet & Mary Beckman. (1988). *Japanese Tone Structure*. Cambridge Massachusetts: MIT Press.
- Piggott, Glyn L. (1987). On the autonomy of the feature nasal. In Bosch, Need & Schiller (eds.) *23rd Annual Regional Meeting of the Chicago Linguistics Society. Part Two: Parasession on Autosegmental and Metrical Phonology*. Chicago Linguistics Society. 223–238.
- Pike, Kenneth L. & E. V. Pike. (1947). Immediate constituents of Mazateco syllables. *International Journal of American Linguistics* 13, 78–91.
- Pollard, Carl & Ivan A. Sag. (1987). *Information-Based Syntax and Semantics* Volume 1, Fundamentals. Center for the Study of Language and Information, Lecture Notes Number 13.
- Poser, William. (1982). Phonological representation and action-at-a-distance. In van der Hulst & Smith (eds.) *The Structure of Phonological Representations*, Part II. Dordrecht: Foris.
- Poser, William. (1989). The metrical foot in Diyari. *Phonology* 6, 117–148.
- Pulleyblank, Douglas. (1986). *Tone in Lexical Phonology*. Dordrecht: Reidel.
- Pullum, Geoffrey K. (1976). *Rule Interaction and the Organization of a Grammar*. New York: Garland.
- Pullum, Geoffrey K. (1989). Topic ... Comment: Formal linguistics meets the boojum. *Natural Language & Linguistic Theory* 7, 137–143.
- Pullum, Geoffrey K. & William A. Ladusaw. (1986). *Phonetic Symbol Guide*. Chicago & London: University of Chicago Press.
- Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence* 13, 81–132.
- Rubin, Philip & Thomas Baer. (1981). An articulatory synthesizer for perceptual research. *Journal of the Acoustical Society of America* 70, 321–328.
- Ryder, Mary E. (1987). An autosegmental treatment of nasal assimilation to labiovelars. In Bosch, Need & Schiller (eds.) *23rd Annual Regional Meeting of the Chicago*

Linguistics Society. Part Two: Parasession on Autosegmental and Metrical Phonology.
Chicago Linguistics Society.

Ryle, Gilbert. (1949). *The Concept of Mind*. London: Hutchinson.

Sagey, Elizabeth. (1986). *The Representation of Features and Relations in Non-Linear Phonology*. Ph.D. Dissertation, Massachusetts Institute of Technology.

Sagey, Elizabeth. (1987). *Phonological Feature Hierarchy*. (Revision of Sagey 1986). Ms, University of California, Irvine.

Sagey, Elizabeth. (1988). On the ill-formedness of crossing association lines. *Linguistic Inquiry* 19, 109–118.

Schane, Sanford A. (1973). *Generative Phonology*. Englewood Cliffs, New Jersey: Prentice Hall.

Schein, Barry & Donca Steriade. (1986). On geminates. *Linguistic Inquiry* 17, 691–744.

Schmiedel, Albrecht. (1988). Temporal Constraint Networks. KIT-Report 69. Technical University, Berlin.

Selkirk, Elisabeth O. (1982). The syllable. In van der Hulst & Smith (1982). *The Structure of Phonological Representations*. Part II.

Slowiaczek, Louisa M. & Daniel A. Dinnsen. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13, 325–341.

Shieber, Stuart. (1986). *An Introduction to Unification-Based Approaches to Grammar*. CSLI Lecture Notes, Number 4.

Shieber, Stuart, Hans Uszkoreit, Fernando C. N. Pereira, Jane J. Robinson, Mabrys Tyson. (1983). The Formalism and Implementation of PATR-II. In Barbara Grosz & Mark E. Stickel (eds.) *Research on Interactive Acquisition and Use of Knowledge* SRI International. 39–79.

Smolka, Gert. (to appear). Feature constraint logics for unification grammars. *Journal of Logic Programming*. Also appeared as IWBS Report 93, IBM Deutschland, November 1989.

Stanley, Richard. (1967). Redundancy rules in phonology. *Language* 43, 393–436.

- Suppes, Patrick. (1972). *Axiomatic Set Theory*. New York: Dover Publications, Inc.
- Swadesh, Morris. (1934). The phonemic principle. *Language* 10, 117–129. Reprinted in Joos (ed.) *Readings in Linguistics I*. Fourth Edition. 32–37.
- Taylor, Barry. (1977). Tense and Continuity. *Linguistics and Philosophy*. 1, 199–220.
- Thráinsson, Höskuldur. (1978). On the phonology of Icelandic preaspiration. *Nordic Journal of Linguistics* 1, 3–54.
- Vennemann, Theo. (1972). Phonological uniqueness in natural generative grammar. *Glossa* 6, 105–116.
- Vincent, Nigel. (1986). Constituency and syllable structure. In Durand (ed.) *Dependency and Non-Linear Phonology*, 305–318. London: Croom Helm.
- Wang, William. (1967). Phonological features of tone. *International Journal of American Linguistics* 33, 93–105.
- Yip, Moira. (1989a). Contour tones. *Phonology* 6, 149–174.
- Yip, Moira. (1989b). Feature geometry and cooccurrence restrictions. *Phonology* 6, 349–374.

Appendix

Appendix A: Extract from Phonological Events¹

1 Introduction

One of the major innovations within post-SPE generative phonology has been the development of frameworks where phonological units are organised in a non-linear fashion. Taking autosegmental phonology (Goldsmith 1976) as our main exemplar of such frameworks, we wish to address the following question: What is the appropriate interpretation of autosegmental representations? There is, of course, a further question about what we mean by *interpretation*: formal, phonetic or computational interpretation? Although we will concentrate on the first of these, we believe that all three aspects should be regarded as closely inter-connected and mutually constraining.

The question of interpreting autosegmental representation has in fact been recently posed by Sagey (1988), and we shall take her proposal as our starting point. While it is uncontroversial to suppose that the relationship between units on a given autosegmental tier is one of temporal precedence, Sagey claims that it is more problematic to pin down what is meant by association between tiers. She argues, cogently we believe, that if association is taken to be a relationship of simultaneity between durationless units, then standard analyses of complex segments and gemination lead to logical inconsistency. Instead, association should be taken as temporal *overlap* between units with duration.

We begin with a review of Sagey's proposals, observing that she adopts an ontology based on points, where intervals are defined as sets of points. We argue that this leads to a number of formal, phonetic, philosophical and cognitive problems, and propose an alternative approach using an ontology based on intervals. In section 2 we define an *event* to be a compound entity consisting of an interval together with a property, and provide axioms governing the overlap and precedence relations which hold between pairs of such events. The resulting ontology, we argue, provides a natural framework within which to model important relationships between phonological gestures. Section 3 (not included here) begins with a presentation of *event*

¹ This section appeared as part of Bird & Klein (1990).

structures, which are collections of events and constraints. We show, with a variety of illustrations, how event structures can be used to formalize the various components of multi-tiered, hierarchical autosegmental representations. We also discuss their close relationship to the notion of a gestural score. It should be stressed at the outset that this article concerns autosegmental representations, and not the rules which are presumed to manipulate them. Due to the expository goals of this paper we have not attempted to carry out a detailed analysis of a large body of phonological data, however we acknowledge that this is an important task and it is one that we intend to undertake in future work.

Deriving the No-Crossing Constraint

Sagey defines three relations on temporal units: simultaneity, precedence and overlap. Certain facts about the first two relations (and presumably the third also) are taken to be 'included in our knowledge of the world' (p.110). We begin with a brief review of these facts.

Temporal overlap is a two-place relation which is reflexive, symmetric and nontransitive. If we employ the notation $x \circ y$ for the statement ' x overlaps y ' then these facts about overlap can be stated as follows:

- (1) a. For any x , $x \circ x$ *overlap is reflexive*
- b. For any x and y , if $x \circ y$ then $y \circ x$ *overlap is symmetric*

If overlap were transitive, a third statement would be necessary:

- (2) For any x , y and z , if $x \circ y$ and $y \circ z$ then $x \circ z$

However, if this were the case we would be back where we began, where association was conceived as simultaneity. Since overlap is nontransitive, we simply omit this statement. (Note that this does not preclude the relation expressed in (2) from holding for a particular choice of x , y and z ; it is just not guaranteed to hold for all such choices.)

Above we described the relation holding between members of a tier as 'temporal precedence'. By this we meant strict linear precedence, which is an irreflexive, asymmetric and transitive relation². We adopt the notation $x \prec y$ to express the statement ' x precedes y ', and write the following expressions (where negation (\neg) is taken to have wider scope than \prec and \circ):

- (3) a. For any x , $\neg x \prec x$ *precedence is irreflexive*
- b. For any x and y , if $x \prec y$ then $\neg y \prec x$ *precedence is asymmetric*
- c. For any x , y and z , if $x \prec y$ and $y \prec z$ then $x \prec z$ *precedence is transitive*

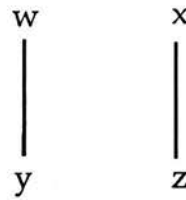
²In fact, these properties only give us a strict partial ordering; to get a linear ordering, we also need an additional statement of connectedness: For all x and y , either x precedes y or $x = y$ or y precedes x . We will return to this once tiers and melodies have been defined. Note that Sagey (1988:110) uses the term 'antisymmetric' when 'asymmetric' is intended (see, for example, Suppes 1972:69 for definitions of these properties).

Perhaps surprisingly, the properties expressed above about overlap and precedence are inadequate in a crucial way. Consider the statement: ' $x \prec y$ and $x \circ y$ '. Clearly, we want this to be inconsistent, given the intended interpretations of \prec and \circ . However, we cannot demonstrate this from what we have said so far. Thus, to express the mutual exclusiveness of overlap and precedence, a further statement is necessary:

- (4) For any x and y , if $x \prec y$ then $\neg x \circ y$

At this point, we seem to have enough machinery to interpret an autosegmental diagram such as (5).

(5)



A line which connects two points, say those labeled w and y , is interpreted as claiming that there is an overlap relation holding between events w and y , while horizontal alignment of two points on the page, say w appearing to the left of x , is interpreted as claiming that a relation of precedence holds between w and x . That is, (5) depicts a situation which we can describe in our notation as follows:

- (6) $w \prec x, y \prec z, w \circ y$ and $x \circ z$.

Now let us consider the situation shown in (7), where two association lines cross:

(7)



We interpret this as shown in (8):

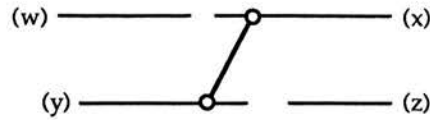
- (8) For some w, x, y and z , (i) $w \prec x$ and $y \prec z$ and (ii) $w \circ z$ and $x \circ y$

However, none of the above facts about overlap and precedence rule out (8) as ill-formed, and a further statement about the relationship between overlap and precedence is therefore necessary. This is given in (9).

- (9) For any w, x, y, z , if $w \prec x$ and $y \prec z$ and $x \circ y$ then $w \prec z$.

In order to help visualise the constraint that is imposed in (9), we adopt Sagey's graphical conventions for representing intervals as labeled time-line segments:

(10)



Given the statements in (4) and (9) about the relationship between overlap and precedence, the no-crossing constraint can be derived. Suppose we have $w \prec x$ and $y \prec z$. From (9) we know that $w \prec z$ and then from (4) that $\neg w \circ z$. The key point here is that the no-crossing constraint does not follow from the definitions of overlap and precedence alone, but from additional statements about their interrelationship.

One apparent virtue of Sagey's approach is that she does not need to stipulate these additional properties. In fact, *all* of the properties of overlap, precedence and their inter-relationship stated above follow from Sagey's conception of intervals as a collection of points. The definitions in (11) are revised versions of those given by Sagey, who employs the notation 'All $P(x)$ ' to refer to the collection of points in an interval x and 'Some $P(x)$ ' to refer to a particular point x . Given her view of an interval as a set of points, there is clearly no distinction to be drawn between x and 'All $P(x)$ '. Moreover, the precise meaning of ' $P(x)$ ' is unclear. Instead we use standard set-theoretical notation to talk about the elements of a set, and formulate the following definitions:

- (11) a. For two intervals x and y , we write $x \prec y$ iff for all $p \in x$ and for all $q \in y$, $p < q$.
- b. For two intervals x and y , we write $x \circ y$ iff for some $p \in x$ and for some $q \in y$, $p = q$.

Given that $<$ is a strict linear ordering on points, it is a relatively straightforward matter to show that (1), (3), (4) and (9) above follow from these definitions and therefore do not require independent statement. This would appear to be a desirable state of affairs, given the economy of statement and simplicity of (11). However, defining intervals in terms of points is questionable from both a philosophical and a cognitive viewpoint.

Although it is a deeply rooted part of our current scientific outlook to regard time as being composed of instants and collections of instants, it has nevertheless been argued by philosophers such as Russell that viewing time as consisting of extended periods which admit ever-finer subdivisions is closer to our pretheoretic intuitions. From a cognitive standpoint, the definitions in (11) are also rather implausible. They suggest that in order for an agent to verify a statement of precedence between two intervals containing an infinity of points, she would have spend forever comparing the points in a pairwise manner; a similarly non-terminating procedure would be required to falsify a statement of overlap between two intervals. Even if one argued that such intervals contained only a finite number of points, the cognitive processing required would be dependent upon the size of the intervals. This is contrary to the seemingly uncontroversial claim that it should take constant time to judge the precedence or overlap of arbitrarily sized intervals.

The above definitions could be rescued from this criticism by referring to the *end-points* of intervals. Let us assume, as before, that $<$ is a strict linear ordering on points, and also that ' $\max(x)$ ' denotes the maximal (i.e. last) element of interval x with respect to $<$, and that ' $\min(x)$ ' denotes the minimal (i.e. first) element of x .

- (12) a. $x < y$ iff $\max(x) < \min(y)$
 b. $x \circ y$ iff $\max(x) > \min(y)$ and $\max(y) > \min(x)$

For the interval endpoints to be specifiable in a way that is independent of the size of the interval (i.e. the number of points it contains, whether finite or infinite), they must be basic to the definition of the interval. In other words, the interval must be defined in terms of its endpoints, rather than as the set of points it contains—for example, as $\{t \mid 3.42 \leq t \leq 3.96\}$, where the numbers represent seconds since the beginning of the utterance. Given the endpoints, it is then a simple matter to determine whether a point is contained in the interval.

However, this position runs into a number of difficulties. First, it is usually difficult to assign a determinate boundary (either perceptually or instrumentally) to the phonetic instantiation of a phonological event. We can be certain about the 'central area' of, say, an interval of nasality or friction in an utterance, but as we near either extremity of such an interval it becomes less certain whether or not a particular point is included in the interval. Consequently, it would seem desirable to allow for a degree of indeterminacy in the location of interval endpoints.

Moreover, even if it were possible from a phonetic point of view to demarcate precisely the beginning and endpoints of some particular event such as voicing, it is highly implausible that one would want to treat such boundaries as part of the phonological specification of a feature or autosegment. This is partially acknowledged by Sagey's claim that "the points of time within a feature or x-slot are accessible only at the late level of phonetic implementation, . . . , they are not manipulable or accessible by phonological rules" (1986:294). Yet (as also pointed out by Hammond 1988:323) this is difficult to reconcile with the fact that points are fundamental to Sagey's ontology.

A related issue is that the phonetic properties of a given point can only be specified in terms of an interval (possibly very small) which contains that point. Thus on Sagey's approach, one first has to construct intervals from points, and only then attach certain properties to these intervals, a two-stage process. This situation is necessitated by a further fact. If a feature is simply considered to be an interval and nothing more, then we could not adequately accommodate a situation where two *distinct* features occupy one and the same interval, because they would then be indistinguishable.

These problems do not arise if intervals are taken as basic to phonology. We believe that Sagey's proposals represent a big step in this direction, but that they do not go far enough. It is perhaps interesting to note that the ontological shift from points to intervals is not new to linguistics; for example, a similar move was made in linguistic semantics by Bennett & Partee (1972).

Recall that intervals have properties attached to them. From now on an interval and a property will be regarded as two aspects of a single unit. When they are bundled together in this way, the result is usually referred to as an *event* (cf. van Benthem

2 Events

The properties of overlap and precedence stated in the last section are gathered together in (13) below. However, from now on, we use the variables w , x , y and z to refer to events rather than intervals. Note that it is unnecessary, and somewhat misleading, to portray events as labeled time-line segments. Events are basic entities in our ontology, having no internal structure other than a particular stated property, and thus can be represented quite adequately by points in our diagrams. As we saw in the preceding section, we can adopt interpretive conventions for standard autosegmental notation whereby association lines correspond to temporal overlap, and left-to-right arrangement on the page corresponds to temporal precedence. Consequently, phonologists can use the usual graphical notation for autosegments and association, while still maintaining the view (if they wish) that autosegments have internal duration.

We said above that an event has a property. This property will correspond to a feature or a gesture. The notion of gesture that we have in mind corresponds broadly to that found in Browman & Goldstein (1986, 1989), Ewen (1986), Lass (1984), Pierrehumbert & Beckman (1988). The latter state '[the elements] could be tones or phonemes, but also demisyllables, articulatory commands, or whatever' (153). Further questions can obviously be raised as to whether such properties play a contrastive role in a phonological system, or whether they are the phonetic realizations of phonological properties. Despite the fact that this is a central issue in developing a detailed theory of event-based phonology, it is one that we cannot address adequately within the confines of this paper, and will therefore sidestep. Our terminology 'phonological events' is intended to be neutral with respect to the phonology/phonetics distinction.

2.1 Axioms for Events

Summarizing from §1, we have the following collection of statements governing a set E of phonological events³:

- (13) a. For any event $x \in E$, $x \circ x$.
Overlap is reflexive (every event overlaps itself).
- b. For any events $x, y \in E$, if $x \circ y$ then $y \circ x$.
Overlap is symmetric (overlapping an event implies being overlapped by it).
- c. For any events $x, y \in E$, if $x \prec y$ then $\neg y \prec x$.
Precedence is asymmetric (preceding an event implies not being preceded by it).
- d. For any events $x, y \in E$, if $x \prec y$ then $\neg x \circ y$.
Precedence is disjoint from overlap (preceding an event implies not overlapping it).
- e. For any events w, x, y and $z \in E$, if $w \prec x$, $x \circ y$, and $y \prec z$ then $w \prec z$.
If one event precedes a member of an overlapping pair of events, and a second event follows the other member of that pair, then the first event precedes the second.

³We are indebted to one of the anonymous *JL* referees for suggesting the use of these English paraphrases. The axioms in (13) and the inference rule in (14) are intended to be couched in a theory of classical first-order logic. We refrain from presenting the proof theory here.

This collection of statements is *minimal*, in the sense that none can be inferred from any combination of the others, and constitute the basic assumptions made. Following standard mathematical practice we will call them *axioms*. Three consequences of these axioms, already discussed in §1, are listed in the Appendix.

We will also presume the following rule of inference:

- (14) **Modus Ponens:** Given a proposition A , and the conditional expression 'if A then B ', infer B .

The reasons for stating axioms are numerous. For example, given a collection of events and certain information about the overlap and precedence relations existing between various units, it is possible to deduce further information. Thus, if we know that a segment p precedes a segment i and that i precedes n , we can infer that p precedes n using (13e), equating x and y . Put slightly differently, the composite statement ' $p \prec i, i \prec n$ and $p \prec n$ ' contains redundant information, given the transitivity of the relation \prec , and so we can abbreviate it by omitting ' $p \prec n$ '.

In addition, it is possible to tell if a set of overlap and precedence statements is consistent; we use the axioms and the inference rule to derive all that can be derived and check that no contradictory statements are present. To summarise then, writing expressions of this kind admits *inference*, *abbreviation* and *consistency checking*. The axiomatic approach is not new to phonology, and has been explored by such linguists as Bloomfield (1926), Bloch (1948), Greenberg (1959) and Batóg (1967). However, there have been few attempts to axiomatize autosegmental phonology.

2.2 Defining Inclusion

Now that axioms for overlap have been provided, it is possible to define temporal inclusion. In fact, inclusion and overlap are interdefinable (van Benthem 1988:59), as shown in the following definitions. (The statement ' $x \subseteq y$ ' should be read: x is included in y .)

- (15) a. For all $x, y \in E$, $x \subseteq y$ iff every $z \in E$ which overlaps x also overlaps y .
 b. For all $x, y \in E$, $x \circ y$ iff there is a $z \in E$ which is included in both x and y .

From this it follows that inclusion is a reflexive and transitive relation. What about symmetry? In general, it will be the case that if $x \subseteq y$ then $\neg x \supseteq y$ (i.e. y is not included in x). Nevertheless, we want to allow the possibility that both $x \subseteq y$ and $x \supseteq y$ hold, corresponding to our intuitive notion of simultaneity. The abbreviatory notation we adopt here is ' $x \equiv y$ ', to be read: x and y are simultaneous (or coterminous). Of course, two events can be simultaneous without being identical, so we favour the use of \equiv over $=$, which Sagey rightly adopts for points and intervals. Here, then, is the definition of simultaneity:

- (16) For all $x, y \in E$, $x \equiv y$ iff $x \subseteq y$ and $x \supseteq y$

A direct consequence is that simultaneity is an equivalence relation, that is, reflexive, symmetric and transitive.

Although we will be mainly concerned with exploring the interpretation of association as overlap, we expect that the interpretation of association as inclusion or as

simultaneity will be useful on occasion, particularly in those cases where the transitivity property is required (e.g. when two autosegments on distinct tiers, linked to the same x-slot, are interpreted as co-articulated, or where association is used to encode hierarchy and 'feature percolation', Clements 1985:250). The inclusion relation may also be useful to express constraints on the spreading of autosegments: if $x \subseteq y$ then x cannot 'spread' beyond the limits of y .

2.3 Homogeneity & Convexity

Now that we have introduced the notion of inclusion, we can ask about the subevents which might be included within a given event. Take, for example, a [+nasal] event e . It is plausible to suppose that all the phonologically relevant subevents of e also have the property of being nasal; that is, the property of nasality is uniformly spread over the whole of e . In this case, we say that the event is *homogeneous*.

By contrast, we might want to claim that a [+stop] event e can be further analyzed as a [+closure] event e_1 followed by a [+release] event e_2 . We can now do so, with a statement to the effect that e includes both e_1 and e_2 . Events which contain distinct subparts in this way will be termed *heterogeneous*⁴.

A related issue arises when we consider a phenomenon such as vowel harmony. We would like to be able to say that the distinctive features common to all of the harmonizing vowels come from a single source, namely the properties of a single event e which overlaps each vowel slot. However, let us consider a sequence $V C V$, where the two V s harmonize, say, for the feature [+back]. This means, in particular, that a [+back] event overlaps both of the V events. Does it also overlap C ? At a phonological level, we do not want to be committed to such a consequence (although it is one which follows on Sagey's account⁵), since the feature [+back] might be either inappropriate or false for the C .

More generally, we are concerned here with a characteristic of events which has been termed 'convexity'. An event e is *convex*, by definition, if it satisfies the following condition:

- (17) For all x_1, x_2, x_3 , if $x_1 \prec x_2, x_2 \prec x_3, x_1 \circ e$, and $x_3 \circ e$ then $x_2 \circ e$.

That is, if e overlaps two events x_1 and x_3 , then it also overlaps any other event x_2 which intervenes between x_1 and x_3 .

An event not satisfying this condition it will be called *non-convex*. We will admit into our framework events of both sorts. Thus, the harmonizing V s in our immediately preceding example will be part of a non-convex [+back] event. This dichotomy allows the local/long-distance spreading distinction (e.g. Hoberman 1988) to be represented.

⁴It has often been observed that what we have called the homogeneous/heterogeneous distinction for events has a parallel in mass/count distinction for objects (e.g. Taylor 1977:210-11).

⁵To see why this is so, consider four intervals x_1, x_2, x_3 and e , where $x_1 \prec x_2, x_2 \prec x_3, x_1 \circ e$, and $x_3 \circ e$. From (11b) we know there are points $p_1 \in x_1$ and $q_1 \in e$ such that $p_1 = q_1$. Now from (11a), for any $p_2 \in x_2, q_1 < p_2$. By a similar argument, for some $q_3 \in x_3, p_2 < q_3$. Therefore $p_2 \in e$, and so $x_2 \circ e$.

Appendix B: Feature Structures and Indices⁶

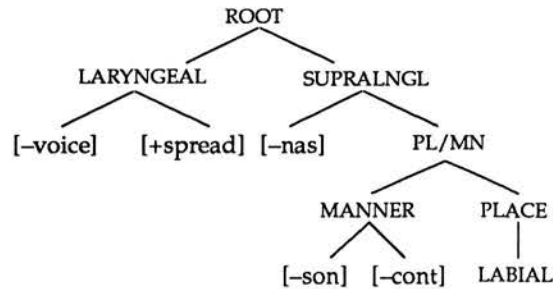
Hayes' (1990) article *Diphthongization and Coindexing* proposes the adoption of a new notation for representing hierarchical phonological structure, in response to a notational problem which he calls the *Diphthongization Paradox*. More generally, he seeks to find a replacement for the current method of depicting feature trees, which is not ideally suited to the clear expression of rules and derivations. While I am in complete agreement with this standpoint and fully support Hayes' assertion that certain ambiguities of interpretation have crept into the graphical conventions currently in widespread use, I nevertheless believe his proposal suffers from a number of problems. Once identified, these problems can be solved with further notational and formal refinements.

Overview. Hayes' main argument goes as follows: if a geminate segment is represented as a single feature tree linked to two C slots, then it is not possible to perform alterations that affect only the first 'half' of the geminate, as required for an analysis of Icelandic preaspiration⁷ and an impressively broad range of other data he provides. The solution, Hayes argues, lies fundamentally in resolving an ambiguity in a graphical notation of phonology, where lines can indicate both association and category membership (cf. Bird & Klein 1990:46–8), illustrated in (18) (Hayes' (13)).

- (18) a. Association Lines: $\begin{array}{c} \text{H} \\ | \\ \text{a} \end{array}$ b. Category Membership Lines: $\begin{array}{c} \sigma \\ \swarrow \downarrow \searrow \\ \text{t} \quad \text{a} \quad \text{p} \end{array} = [\text{tap}]_{\sigma}$

In order to resolve this ambiguity, the distinction must be signalled explicitly: two notational devices are required. The first is for category membership or what I will henceforth call *dominance*. Hayes represents the feature geometry of (19a) as (19b).

- (19) a.



⁶ This section will appear as Bird (1991a).

⁷ Note that Hayes' analysis of Icelandic aspiration (along with those of Clements (1985), Sagey (1986) and others) is based on an assumption that a lengthening rule exists which feeds the preaspiration rule for forms like /opna/ → /oppna/ → [chpna]. However, Árnason (1986:13) claims that the evidence for such a rule is unclear. Furthermore he argues that stress data is most naturally accounted for if the aspiration forms part of the syllable nucleus (as a voiceless vowel). I suspect, therefore, that an adequate account of this data will be less straightforward than Hayes' rule (34) suggests.

- b. ROOT: LARYNGEAL: [–voice]
 [+spread]
 SUPRALNGL: [–nas]
 PL/MN: MANNER: [–son]
 [–cont]
 PLACE: LABIAL

Note that (19b) is essentially the same notation that Gazdar et al. (1985:45) used for syntax trees. Another variant is known as the *attribute-value matrix* or the *feature structure* (e.g. Kaplan & Bresnan 1982, Shieber 1986, Broe *to appear*).

Although feature structures like (19b) look rather different to trees, Hayes stresses that the difference is only superficial and the importance of the new notation to phonology lies in any added convenience that may result from its adoption. However, there is a further advantage to the use of this notation, because it has been given a rigorous mathematical foundation (e.g. Johnson 1988); a similar foundation for the purely graphical notations of phonology is non-existent. Diagrams are often the most apt way to represent information (Larkin & Simon 1987), but coupled with their flexibility is an inherent difficulty for formalization.

The second notational device is for association. Hayes proposes to annotate the nodes in these trees with indices, where coindexation represents association (an idea attributed to Halle & Vergnaud 1980). Hayes' reformulation of the No Crossing Constraint ensures that the indices can be given a temporal ordering.

Hayes' two notations are combined as follows. Each member of the skeletal tier is given a single index. This sequence serves as a series of reference points that do not appear to be subject to alteration. A 'percolation convention' causes an index to be copied to the corresponding root node and then downwards through the entire feature tree, as illustrated in (20) (Hayes 1990:44).

- (20) a. *Percolation Convention:*

When indices are assigned to or removed from a node N , the assignments and deletions are automatically carried over to all nodes dominated by N .

- b. /õ:/ : V₁ V₂
R₁₂ : L: [+voice]
S: [+nas]
PM: M: [+cont]
[-cons]
P: LB: [+round]
D: [-high]
[-low]

c.	V ₁	V ₂	
	R ₁₂ :	L ₁₂ :	[+voice] ₁₂
		S ₁₂ :	[+nas] ₁₂
		PM ₁₂ :	M ₁₂ :
			[+cont] ₁₂
			[-cons] ₁₂
		P ₁₂ :	LB ₁₂ :
			[+round] ₁₂
		D ₁₂ :	[-high] ₁₂
			[-low] ₁₂

As for long vowels, geminates are viewed as a single structure dominated by two adjacent skeletal tier elements. All nodes of a geminate therefore bear two indices. Crucially, this gives a handle on the 'two halves' of a geminate at all levels of hierarchy, thereby making it possible for a rule to affect only one half of a geminate. In the case of Icelandic preaspiration the diphthongization rule simply deletes the 'first' of the two indices from all nodes dominated by (and including) the place/manner node, thus it "delinks the place/manner autosegment".

Discussion. Three aspects of Hayes' proposal seem problematic. The first concerns the effect of the rule of autosegment delinking described above. If rules are able to remove indices then it is possible, in principle, for a node to lack an index. Is such a node floating, or has it effectively been deleted? Indeed, is this distinction preserved in Hayes' proposal? Suppose that an indexless node is to be interpreted as floating; it may later gain an index and effectively be 'relinked'. Under this view we can think of indices as 'alignment constraints'. An indexless node is unconstrained as to its structural alignment. Those nodes remaining indexless at the end of a derivation could be construed as phonetically uninterpretable and effectively deleted (cf. stray erasure). The other interpretation of indexless nodes is that they have been deleted. This is the option which Hayes (pers. comm.) has advocated. His suggestion is as follows (the bracketed part is my addition):

"Rules of the form $i \rightarrow \emptyset$ [where i is the sole index of a node] are deletion rules (i.e. the node on the target tier indexed i is deleted, along with the nodes it dominates, owing to percolation). On the other hand, rules of the form $i \rightarrow i'$, where i' represents an index not previously present in the representation, could be taken to be rules of delinking."

This proposal could probably be made to work if the following apparent problems are dealt with. First, consider the situation where a node x_2 dominates a node y_{23} . A rule $2 \rightarrow \emptyset$ would delete the index from x_2 , thereby deleting the node itself. According to Hayes' definition above, since x_2 dominates y_{23} , y_{23} would also be deleted even though it bears another index. However, as y_{23} is effectively doubly associated, its deletion would seem to be an undesirable consequence of the deletion of x_2 . Second, it seems necessary to adopt a convention that any rule assigning an index to a node must also remove an existing index just in case the existing index is unique. Therefore, some revision of the above definition appears necessary.

We have now seen two possible solutions to the problem of indexless nodes and delinking versus deletion which arise from Hayes' (1990) proposal. One solution (due to Hayes) is to view indexless nodes as having effectively been deleted, and to

adorn floating nodes with unique indices. Another solution—the one I favour—is to view indexless nodes as floating (or delinked) and for delinked nodes to be deleted at the end of a derivation by stray erasure.

These notions of delinking and deletion have further ramifications. If, at some stage in a derivation all of the '2' indices have been removed, then the adjacency of certain nodes (say y_1 and y_3) would not be naturally represented. Hayes states (p. 45) that "indices in rules are meant to be consecutive". Therefore, a rule such as the one for assimilation (Hayes' (24)) cannot apply, even though y_1 and y_3 are now adjacent. One solution would be to re-index nodes so that an index 3 is changed to 2, a 4 to 3, and so on throughout the entire structure. A more attractive solution, I believe, would involve the use of abstract indices, say $\iota_1, \iota_2, \iota_3, \dots$ and an ordering relation $<$ on these indices. Two nodes could be said to be adjacent just in case there is no other node with an intervening index. (See Bird & Klein 1990:42 for a more detailed discussion of adjacency.)

The second problematic aspect of Hayes' proposal concerns the representation he proposes for contour segments (p. 60), where "the two feature values receive the same index, with their temporal ordering determined simply by their ordering within the representation". Accordingly, Hayes' example (55) for /č/ includes a specification for the manner tier as in (21).

(21) $M_1 : [-\text{cont}]_1 [+ \text{cont}]_1$

This leads to a contradiction. We are told that coindexing encodes association (p. 43), and that associated elements are "pronounced together" (p. 40). Therefore $[-\text{cont}]_1$ and $[+ \text{cont}]_1$ must be pronounced together, which is impossible. The very same mechanism which replaces association lines will autosegmentally 'link' the halves of an affricate⁸. The result is an instance of Sagey's paradox (Sagey 1986:282ff). Hayes' proposal might be rescued from this contradiction by appealing to a 'context-sensitive' interpretation of indexing. For example, we could say that coindexation only encodes coarticulation across tiers, not within tiers⁹. However this option is ruled out by the later claim that "only the indexing is formally significant" (p. 63). This claim requires that indexing is interpreted independently of any other aspects of notation (such as the tier structure), and so a contradiction is inevitable.

The heart of the problem is that Hayes' indices correspond to *atomic* intervals which, by definition, cannot be subdivided. One solution, following Bird & Klein (1990), is to permit the ever-finer subdivision of intervals. Then the manner tier specification could be $M_1 : [-\text{cont}]_{1a} [+ \text{cont}]_{1b}$, where 1a and 1b are the two halves of 1, just as 1 and 2 are the halves of 12. This provides a natural expression of the adjacency of 1b and 2 and the non-adjacency of 1a and 2.

The final problem concerns the *spurious ambiguity* noted in Section 9.4, where (22a) and (22b) (simplifying from Hayes' example (59)) are "phonologically distinct, though phonetically identical".

⁸ I am grateful to Bruce Hayes for suggesting this sentence as a paraphrase of my argument.

⁹ Note that the use of consecutive indices to encode adjacency would need revision for the edge effects claimed for affricates (Sagey 1986:93ff) to be expressed.

(22) a. C_1 C_2 Output: [bm]

$PL/MN_1 :$ $MANNER_1 :$ $[-son]_1$ $[-cont]_1$ $PLACE_{12} :$ $LABIAL_{12}$	$PL/MN_2 :$ $MANNER_2 :$ $[+son]_2$ $[-cont]_2$	
--	--	--

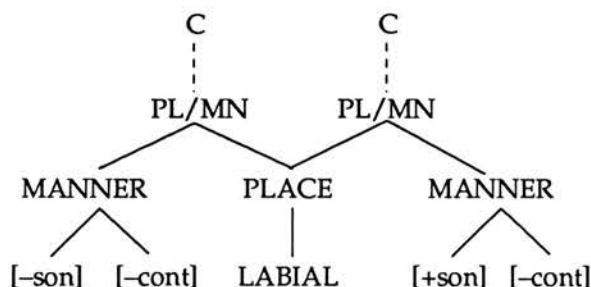
b. C_1 C_2 Output: [bm]

$PL/MN_1 :$ $MANNER_1 :$ $[-son]_1$ $[-cont]_1$	$PL/MN_2 :$ $MANNER_2 :$ $[+son]_2$ $[-cont]_2$ $PLACE_{12} :$ $LABIAL_{12}$	
--	---	--

Hayes' solution involves revising the definition of dominance and the percolation convention, which apparently has the effect that "the putatively distinct outputs in (59) are notational variants. That is, while we may arrange nodes in outline form for convenience, only the indexing is formally significant." However, if only the indexing were formally significant then the percolation convention (p. 44, revised on pp. 63–4) could not operate, since it requires access to information about dominance. In short, Hayes has (i) defined a property P (i.e. dominance), then (ii) shown how a property Q (i.e. coindexing) is derived from P, and then (iii) claimed that P is unnecessary. Crucially, he has failed to show how Q may be derived in the absence of P. Therefore step (iii) is in doubt.

A more attractive solution to this spurious ambiguity problem is to permit C_1 and C_2 to *share* their place of articulation. Example (23a) illustrates the solution (the dashed lines indicate elided structure). The corresponding feature structure appears in (23b), a notation discussed by Shieber (1986:13) and Johnson (1988:17–18).

(23) a.



b. C_1 C_2 Output: [bm]

$PL/MN_1 :$ $MANNER_1 :$ $[-son]_1$ $[-cont]_1$ $PLACE_1 :$ $LABIAL_{12}$ 1	$PL/MN_2 :$ $MANNER_2 :$ $[+son]_2$ $[-cont]_2$ $PLACE_2 :$ 1	
---	---	--

In (23b), both instances of 1 are interpreted as referring to the one place node: both C_1 and C_2 effectively share the information that the place of articulation is labial.

In conclusion, Hayes' solution to the diphthongization paradox relies on clearly distinguishing dominance and association: "once we separate the two functions, the paradox disappears" (p. 40). However, his percolation convention (21)—loosely:

dominance implies association—and his rule (60)—loosely: association implies dominance—leave the reader in doubt that he has actually achieved a clear distinction between dominance and association. I agree with Hayes' diagnosis but not with his solution, which appears to suffer from a number of flaws. I have indicated how these flaws might be dealt with, drawing on insights discussed at greater length in (Bird & Klein 1990, Bird 1990). My proposals may be summarized as follows: (i) the adoption of abstract indices related by a temporal ordering (for the deletion and delinking problems), (ii) the use of 'ever-finer subdivision', where indices are thought of as temporal intervals which may be partitioned (for the problem with affricates), and (iii) the use of structure-sharing (for the spurious ambiguity problem).

A final comment about the ultimate purpose of these proposals is in order. Hayes is not alone in his frequent use of the term 'formal' (and the related terms 'formalize' and 'formalism'). However, far from being a well-defined notion as might be expected, the word is an enigma (cf. Pullum 1989). Hayes' theory is formalized simply by formulating a notation (pp. 43–4). I contend that the more substantive notion of 'formal' as it appears in logic should also be adopted in phonology. A grammar would only be called formal once it was expressed in a descriptive language having both a formal syntax and a formal semantics. (To be sure, there are other properties we would like a grammar to have, such as convenience, clarity and economy. The requirement for a formal syntax and a formal semantics is proposed only as a necessary condition and not also as a sufficient condition for the acceptability of a grammar.) Once this *formal adequacy* has been achieved, the empirical adequacy of a grammar that uses feature-structures can be tested automatically. There exist software systems such as PATR-II (Shieber et al. 1983) which can mechanically grind out the empirical consequences of feature-structure based linguistic theories. The above formal refinements to Hayes' theory are crucial, I believe, if the coverage of a large grammar couched in his theory is to be determined.